

Pro gradu -tutkielma

**Tulevaisuuden työkalu metsien suojeluarvon
määrittämiseen? – Boreaalisten puulajien tunnistus
hyperspektrikuvauksen avulla**

Laura Kauniskangas



Jyväskylän yliopisto

Bio- ja ympäristötieteiden laitos

Ympäristötiede

9.6.2022

JYVÄSKYLÄN YLIOPISTO, Matemaattis-luonnontieteellinen tiedekunta
Bio- ja ympäristötieteiden laitos
Ympäristötiede

Laura Kauniskangas: Tulevaisuuden työkalu metsien suojelevarvon määritykseen? - Boreaalisten puulajien tunnistus hyperspektrikuvauksen avulla
Pro gradu -tutkielma: 57 s.
Työn ohjaajat: Yliopistonlehtori Anssi Lensu ja FT Pauliina Salmi
Tarkastajat: Yliopistonlehtori Elisa Vallius ja FT Pauliina Salmi

Kesäkuu 2022

Hakusanat: drone, FPI-kamera, hyperspektrikamera, kaukokartoitus, luokittelu, ohjattu koneoppiminen, puulajintunnistus, RPAS

Puulajien tunnistusta hyperspektrikuvista, yhdistettynä nykyiseen metsäinventointiin, pidetään potentiaalisena keinona lisätä metsänhoidon kustannustehokkuutta sekä kartoittaa metsien suojelevarpeita kattavasti. Tutkimusta tarvitaan kuitenkin edelleen, jotta löydetäisiin tehokkaimmat kaukokartoitusmenetelmät eri puulajien ja metsätyyppien tutkimiseen. Tämän tutkielman tavoitteena oli kehittää menetelmä, jolla boreaalisen metsän puulajeja voitaisiin tunnistaa hyperspektrikuvista vähintään 75 % tarkkuudella hyödyntäen ohjattua koneoppimista. Tutkimukseen sisältyneet lajit olivat harmaaleppä, koivu, mänty ja kuusi. Aineisto koostui Äänekoskella sijaitsevasta metsäisestä puronvarresta kauko-ohjattavalla ilma-aluksella (RPAS) kuvatuista hyperspektrikuvista. Lisäksi käytössä oli referenssimaastodataa. Hyperspektrikuville tehtiin kalibrointi, pikseleiden kohdistus sekä georeferointi Maanmittauslaitoksen ortokuva-aineistoa käyttäen. Puiden kruunuista poimittiin pikseleitä ja kuvia esikäsiteltiin pikseleiden kirkkauserojen sekä kruunujen välisten valojen ja varjojen tasoittamiseksi R-ohjelmalla. Tämän jälkeen puulajeja luokiteltiin lineaarisen diskriminanttianalyysin (LDA), Random Forestin (RF) ja tukivektorikoneen (SVM) avulla. Korkein luokittelun kokonaistarkkuus 98 % (Kappa = 0,96) saatiin LDA:n avulla mutta kaikilla luokittelijoilla yllettiin yli 77 % kokonaistarkkuuksiin kaikilla testatuilla esikäsitelymenetelmillä. Täten voidaan todeta, että tutkimuksen menetelmät

vaikuttavat soveltuvan hyvin boreaalisten puiden luokitteluun ja tueksi metsien kestävän käytön arviointiin.

UNIVERSITY OF JYVÄSKYLÄ, Faculty of Mathematics and Science
Department of Biological and Environmental Science
Environmental Science

Laura Kauniskangas: Future tool for evaluating protection needs of boreal forests? – Detecting boreal tree species with hyperspectral imaging
MSc thesis: 57 p.
Supervisors: Senior lecturer Anssi Lensu and PhD Pauliina Salmi
Inspectors: Senior lecturer Elisa Vallius and PhD Pauliina Salmi

June 2022

Keywords: drone, FPI sensor, hyperspectral imager, remote sensing, RPAS, supervised learning, tree species classification

According to recent studies hyperspectral imaging shows promise in supporting more traditional forest inventory methods in increasing cost-efficiency of forestry as well as conservation of vulnerable forest areas. However, research still lacks conclusive results about the most suitable remote sensing methods to detect different forest types and tree species. Our intention was to find a method that could classify four boreal tree species with minimum 75% accuracy using supervised classification methods. Hyperspectral images with 55 spectral bands were used. They had been acquired outside of this study from a boreal forest located in Äänekoski (Finland) using a drone. The species of each tree was validated according to GPS field measurements. Half of the species selected were coniferous (Scots pine and Norway spruce) and the other half deciduous (birch and grey alder). Hyperspectral images were calibrated, coregistered, and georeferenced using orthophoto data by National Land Survey of Finland. The pixels of each tree crown were extracted, and images preprocessed in R to even out illumination differences between tree canopy pixels. After this, the images were classified using Linear Discriminant Analysis (LDA), Random Forest (RF) and Support Vector Machines (SVM). The highest overall accuracy achieved was with LDA (98%, Kappa = 0,96) whilst all the classifiers performed well resulting in higher than 77% overall accuracies with every preprocessing method used. To conclude, the methods used

were efficient in discriminating between the four boreal tree species. Thus, they might be useful in supporting the sustainable usage of Northern forests.

SISÄLLYSLUETTELO

1 JOHDANTO.....	9
2 TUTKIMUKSEN TAUSTA	13
2.1 Hyperspektritekniologia, -aineistot ja RPAS-laitteet.....	13
2.2 Boreaalisten puiden spektrisiä ominaisuuksia	16
2.3 Hyperspektrikuvien esikäsittely.....	19
2.4 Hyperspektrikuvien luokittelu ohjatun koneoppimisen menetelmillä	21
2.5 Tutkimustuloksia hyperspektrikuvien käytöstä puulajien luokittelussa	26
3 AINEISTO JA MENETELMÄT	31
3.1 Aineiston keruu, tutkimusalue ja välineet	31
3.2 Aineisto.....	36
3.3 Aineiston esikäsittelymenetelmät.....	37
3.4 Puulajien luokittelumenetelmät.....	42
4 TULOKSET	43
5 TULOSTEN TARKASTELU	47
KIITOKSET	52
KIRJALLISUUS	52

LYHENTEET

CIR	Color-infrared; CIR-kuvat on otettu näkyvän valon ja lähi-infrapunaa aallonpituusalueelta
CNN	Convolutional neural network; Ohjatun oppimisen luokittelija (selitetty kappaleessa 2.4)
DN	Digital Number; pikselin numeerinen intensiteetti-arvo
FPI	Fabry-Pérot-interferometri; FPI-hyperspektrikameran optinen komponentti, jolla määritetään sensorin tallentamat aallonpituuskaistat
k-NN	K-nearest neighbor; Ohjatun oppimisen luokittelija
LDA	Linear discriminant analysis; Ohjatun oppimisen luokittelija (selitetty kappaleessa 2.4)
MLP	Multilayer perceptron; Ohjatun oppimisen neuroverkko-luokittelija (selitetty kappaleessa 2.4)
NIR	Near-infrared; Lähi-infrapunasäteily (noin 750–2500 nm)
NN	Neural network; Ohjatun oppimisen luokittelija (selitetty kappaleessa 2.4)
RF	Random forest; Ohjatun oppimisen luokittelija (selitetty kappaleessa 2.4)
RGB	Red, Green, Blue; aito- eli tosivärikuva, joka sisältää näkyvän valon punaisen, vihreän ja sinisen aallonpituudet
RPAS	Remotely Piloted Aircraft System; Kauko-ohjattu ilma-alus
SVM	Support vector machine eli tukivektorikone; Ohjatun oppimisen luokittelumenetelmä (selitetty kappaleessa 2.4)
UAV	Unmanned Aerial Vehicle; Miehittämätön ilma-alus
SWIR	Short-wave infrared; Lyhytaaltainen infrapunasäteily (noin 1000–2500 nm)

VNIR

Visible and near-infrared; Näkyvän valon ja lähi-infrapunan aallonpituudet (noin 400–1000 nm)

1 JOHDANTO

Hyperspektrikuvauksessa auringonsäteilyn absorptiota ja heijastuksia maastosta tutkitaan runsaasti eri aallonpituuskanavia tallentavalla sensorilla (Alcolea ym. 2020). Hyperspektrikuvausta voidaan käyttää ympäristöntutkimuksessa materiaalien tunnistamiseen. Suomessa kaukokartoituksena tapahtuva puustoinventointi perustuu tällä hetkellä pääosin laserkeilaukseen ja ilmakuvaukseen, joilla ei kuitenkaan saada tilastollisesti tarkkaa tietoa puulajien tiheyksistä (Nevalainen ym. 2017). Puiden ominaisuuksia ja lajeja pystyttäisiin yksilöimään tarkemmin käyttämällä korkeamman spektrisen ja spatiaalisen resoluution kaukokartoituskuvia. Näihin tarkoituksiin hyperspektrikuvausta voidaan pitää varsin soveltuvana menetelmänä. Lisätutkimukselle on kuitenkin edelleen tarvetta, sillä moniulotteisen informaation tutkiminen asettaa omat haasteensa. Korkean resoluution aineistot vaativat yhä tehokkaampia ja tarkempia aineistonkäsittelymenetelmiä (Turner ym. 2014). Tästä huolimatta hyperspektrikuvaus on kovaa vauhtia kehittyvä menetelmä ja sen potentiaaliset hyödyt on tunnistettu esimerkiksi täsmämetsätaloudessa, metsän biomassan estimoinnissa sekä kasvun mallintamisessa (Nevalainen ym. 2017, Näsi ym. 2016). Tuottamalla tietoa puuston lajikoostumuksesta ja -tiheyksistä ymmärrettäisiin paremmin myös metsien monimuotoisuuden tilaa (Saarinen ym. 2017).

Suomen metsien inventointi toteutetaan maastossa tehtävin koealamittauksin noin viiden vuoden välein (Luke 2021). Metsäinventointien avulla tuotetaan tilastollista tietoa muun muassa metsien puulajikoostumuksesta, mikä on tarpeellista kansallisen ja kansainvälisen päätöksenteon tukena, metsätalouden ja -teollisuuden lähdeaineistona sekä metsien käytön kestävyden arvioinnissa. Lisäksi metsien lajikoostumuksen määrittäminen on oleellista ekosysteemien toiminnan, häiriöiden ja biodiversiteetin tutkimuksessa (Dadon ym. 2019, Fricker ym. 2019). Viimeisimpien mittausvuosien aikana valtakunnan metsien inventoinnissa (VMI) on panostettu monilähteiseen inventointiin, jossa maastomittauksiksi yhdistetään kaukokartoitusaineistoista saatavaa numeerista tietoa. Satelliittikuvista

ja vastaavista aineistoista saadaan tarkempia mittaustuloksia pienemmille alueille, esimerkiksi kuntakohtaisesti. Lisäksi metsäalueisiin saadaan liitettyä paikkatietoja. Kaukokartoituksen implementointia osaksi metsien kansallista monitorointia pyritään edistämään kustannusten vähentämiseksi sekä laadukkaampien ja nopeampien mittausten tuottamiseksi (Luke 2015).

Suomea sitoo velvoite raportoida kuuden vuoden välein EU:n direktiiviluontotyyppien määrä ja tila. Suojeltavia Natura 2000 -biotooppeja esiintyy maassamme tällä hetkellä 68 kappaletta, joista 11 on metsätyyppejä (Ymparisto.fi 2020). EU:n taholla ollaan huolestuneita habitaattien tilan heikentymisestä ja tavoitellaan luonnonsuojeludirektiivien tehokkaampaa täytäntöönpanoa (European Environment Agency 2020). Mittavat, koko valtion kattavat luontotyyppien kartoitukset edellyttävät kustannuksellisesti ja ajankäytöllisesti tehokkaita tapoja kerätä aineistoa. Työvoima on kallista ja kenttämittaukset ajallisesti vaativia, joten kaukokartoitusmenetelmien implementointi nykyisiin käytäntöihin voisi tehostaa tiedonkeruuta ja -käsittelyprosessia. Hankkimalla tietoa boreaalisille metsille tyypillisten puulajien spektri-profiileista ja spektrisestä luokiteltavuudesta voidaan kehittää koneoppimisen menetelmiä, joita voisi soveltaa esimerkiksi suojeltujen metsien esiintymisen ja tilan tarkkailuun. Spatiaalisesti tarkkojen puustokarttojen tuottamisesta voisi lisäksi olla hyötyä ekosysteemien hoidossa, ja esimerkiksi kuolleiden puiden identifiointi auttaa luonnon häiriöiden aiheuttamien metsätuhojen arvioinnissa (Fricker ym. 2019).

Nezamin ym. (2020) mukaan lähietäisyydeltä tapahtuva kaukokartoitus tarkoittaisi metsäinventoinnin osalta nopeampia ja edullisempia mittauksia sekä joustavampaa ja kohdistetumpaa aineiston hankintaa ja tietojen käsittelyä manuaalisen työn vähentyessä. Kauko-ohjattavien ilma-alusten (RPAS, Remotely piloted aircraft system) ja hyperspektrisensorien teknologisen kehityksen ansiosta niiden käyttö ympäristöntutkimuksessa on nopeasti yleistynyt. RPAS-laitteisiin voidaan nykyään liittää erittäin tarkkan spektrisen ja spatiaalisen resoluution kuvausinstrumentteja,

mikä mahdollistaa uudenlaisia sovelluskohteita ja parantaa niiden kustannustehokkuutta suhteessa perinteiseen satelliittikaukokartoitukseen.

Hyperspektrikamerat ovat kevyitä, tehokkaita ja niiden saavutettavuus on huomattavasti parantunut. RGB (Red, Green, Blue) -kuvaus on kuitenkin edelleen hyperspektrikuvausta edullisempaa ja siten myös suositumpaa useissa tutkimustarkoituksissa (Nezami ym. 2020). Hyperspektri- ja RPAS-kuvauksen yhdistelmällä saadaan tuotettua RGB-kuvausta moniulotteisempaa ja tarkempaa informaatiota, joka oikeilla menetelmillä eristettynä voisi johtaa entistä täsmällisempiin luokittelutuloksiin puulajien osalta. Tutkijoita ovatkin kiinnostaneet muun muassa RPAS-laitteista otettujen hyperspektrikuvien käytettävyys yksittäisten puiden ja esimerkiksi sairauksista, hyönteisistä ja paloista aiheutuneiden puustotuhojen tunnistukseen sekä metsätalousalueiden ja fenologisten ilmiöiden monitorointiin (Eugenio ym. 2020).

Kuten sanottua, hyperspektrikuvauksen soveltuvuutta metsien lajikoostumuksen määrittämiseen on lähivuosina tutkittu paljon niin Suomessa (mm. Näsi ym. 2016, Nevalainen ym. 2017, Saarinen ym. 2017, Tuominen ym. 2017) kuin ulkomailla (mm. Mozgeris ym. 2018, Trier ym. 2018, Miyoshi ym. 2020). Kirjallisuus ei kuitenkaan vielä vastaa laajojen, tiheiden ja monimuotoisten metsien kuvantamiseen liittyviin haasteisiin (Modzelewska ym. 2020). Rakenteellisesti yksinkertaisempienkin metsien hyperspektrikaukokartoitukseen käytetään yhä monia erilaisia menetelmiä, joiden joukosta voi olla haastavaa valita kuhunkin tutkimustarkoitukseen parhaiten soveltuva menetelmä.

Avaruusjärjestöillä on käynnissä globaaleja ohjelmia, joissa kehitetään hyperspektrikuvausta satelliiteista ympäristön yksityiskohtaisempaa monitorointia varten (Pascucci ym. 2020). Voidaan ajatella, että hypoteettisesti olisi mahdollista saavuttaa tilanne, jossa tiettyjen metsätyyppien puulajeja voisi tunnistaa luotettavasti satelliittiin sijoitettavalla hyperspektrikameralla jopa yksittäisten puiden tarkkuudella. Puulajitunnistuksen tutkimisesta hyperspektrikuvista on tässä apua, sillä olisi löydettävä rajallinen määrä kapeahkoja taajuuskaistoja, joilla

lajit ovat erotettavissa toisistaan. Laaja-alainen ja jatkuva puulajien hyperspektritutkimus satelliiteilla ei kuitenkaan vielä ole mahdollista ja kehitys kohti sen käyttöönottoa edellyttäisi yksittäisten puulajien spektri-profiilien yksityiskohtaista tuntemista. Jos puulajien tunnistukseen tarvittavien taajuuskaistojen määrää onnistuttaisiin vähentämään, mahdollistaisi se suurempien alueiden tutkimisen lajiluokittelua ajatellen.

Tämän tutkielman kirjallisuuskatsauksen perusteella noin 75 % puista tai enemmän on pystytty tunnistamaan hyperspektrikuvien ja muun aineiston avulla aiemmissä tutkimuksissa eli luokittelemaan oikeaan luokkaan ohjatun koneoppimisen menetelmin. Tässä tutkimuksessa pyrittiin selvittämään, ovatko yksittäiset puut luokiteltavissa lajin perusteella vastaavalla tarkkuudella (vähintään 75 %) RPAS-laitteella ja FPI-hyperspektrikameralla (FPI, Fabry-Pérot-interferometri) kerätystä kaukokartoitusaineistosta. Käytetty aineisto oli kerätty boreaaliseen metsästä Äänekosken Pykälisestä. Hyperspektrikuvat sisälsivät spektrikanavia näkyvän valon ja lähi-infrapun aallonpituuksilla (VNIR; 500–900 nm). Aineistoa esikäsiteltiin ja tutkimukseen valituista puista poimittiin pikseleitä niiden latvuksista, joiden luokittelua testattiin ohjatun koneoppimisen menetelmillä. Analyysissä käytetyt luokittelijat olivat LDA, RF ja SVM. Pro gradu -työn tutkimuskysymyksiksi määriteltiin seuraavat:

1. Voidaanko yli 50-taajuuskaistaa sisältävien hyperspektrikuvien avulla tunnistaa havu- ja lehtipuulajeja (mänty, kuusi, harmaaleppä ja koivu) vähintään 75 % varmuudella borealisessa metsässä?
2. Mitkä tutkimukseen valituista esikäsitely- ja luokittelumenetelmistä toimivat parhaiten puulajien tunnistuksessa käytettäessä monitaajuuskaistaista kuvainformaatiota?
3. Parantavatko muut ilmakuvista laskettavissa olevat piirteet, kuten peräkkäisten aallonpituuksien heijastavuuksien erotukset tai kasvillisuusindeksit, kokonaisluokittelutarkkuutta?

2 TUTKIMUKSEN TAUSTA

2.1 Hyperspektritekniologia, -aineistot ja RPAS-laitteet

Kuvausspektrometrit tuottavat kolmiulotteista dataa, joka muodostuu spatiaalisista x- ja y-ulottuvuuksista (koordinaatit maan pinnalla) sekä spektrisestä λ -ulottuvuudesta (Hagen ym. 2012). Niiden toiminta perustuu kuvauspinnan heijastaman sähkömagneettisen säteilyn mittaamiseen. Grahnin ja Geladin (2007) mukaan kuvaussensorin spatiaalisella resoluutiolla viitataan sen tallentamien pikseleiden määrään kuvassa. Mitä enemmän pikseleitä on sitä korkeampi spatiaalinen resoluutio, mikä vaikuttaa kuvattavalla pinnalla olevien, spatiaalisten yksityiskohtien erotettavuuteen. Tähän vaikuttavat myös optiikka ja kuvausetäisyys. Spektrinen resoluutio puolestaan kertoo hyperspektrikuvan koostamisessa käytetystä aallonpituuskaistojen määrästä. Spektrinen resoluutio on korkea ja spektri informatiivinen, kun mittaustuloksia eri taajuuskaistoilta on paljon. Radiometrinen resoluutio viittaa pikseleiden digitaalisiin harmaa-arvoihin eli intensiteettiin. Satelliittikuvat voivat esimerkiksi olla 8- tai 10-bittisiä, jolloin niistä on erotettavissa 256 tai 1024 kirkkausarvoa tai sävyä (Grahn ja Geladi 2007).

Aasenin ym. (2015) mukaan kaukokartoitussensoreita voidaan luokitella aineiston keräämistävän sekä tyyppin perusteella. Aineiston tyyppin mukaan nämä sensorit voidaan luokitella kahteen luokkaan; markkinoille ensin tulleisiin multispektrikameroihin ja niitä seuranneisiin hyperspektrikameroihin. Multispektrikamerat tuottavat radiometrisesti ja spektrisesti karakterisoitua dataa mutta leveillä ja epäjatkuvilla aallonpituuskaistoilla. Niiden spektrinen resoluutio on heikompi kuin hyperspektrikameroiden, jotka tallentavat määrällisesti enemmän jatkuvia ja kapeita aallonpituuskaistoja (Aasen ym. 2015). Tämän lisäksi hyperspektrikameroilla voidaan kattaa laajempi spektrinen alue. Perinteiseen spektrofotometriteknologiaan perustuvan spektrisen informaation lisäksi hyperspektrikamerat tallentavat sijaintitietoja kuvauskohteesta (Grahn ja Geladi 2007). Yksi syy hyperspektrisensoreiden kasvavaan suosioon lähietäisyydeltä

tapahtuvassa kaukokartoituksessa on siis niiden kyky tuottaa aiempaa monipuolisempaa ja tarkempaa informaatiota sekä spatiaalisesti että spektrisesti.

Ensimmäisillä hyperspektrisensoreilla kohdetta mitattiin pisteskannaamalla (engl. whiskbroom) tai riviskannaamalla (engl. pushbroom) mutta nykyisissä lähietäisyydeltä tapahtuvissa kaukokartoitustehtävissä hyödynnetään spatiaalista 2D-informaatiota tuottavia sensoreita. Aineiston keruutavan perusteella kuvauslaitteet voidaan jakaa kahteen eri luokkaan. Kameroita, joilla kanavat tallennetaan eri sensoreilla, eriaikaisesti tai peräkkäisesti, voidaan nimittää *image-frame*-sensoreiksi (Aasen ym. 2015). Tällaisia ovat esimerkiksi FPI-kamerat, joilla eri aallonpituudet tallennetaan peräkkäin etukäteen määritettyyn interferometrin ilma-aukon leveyksien sekvenssiin perustuen. *Snapshot*-kamera sen sijaan tallentaa kaikki kuvan taajuuskaistat samanaikaisesti ilman skannausprosessia. Tällainen on esimerkiksi kolme taajuuskaistaa tallentava, tavallisessa värivalokuvauksessa käytettävä, RGB-kamera (Hagen ym. 2012, Aasen ym. 2015).

Näkyvän valon aallonpituuksia tallentavat hyperspektrisensorit ovat edelleen eniten tutkimuskäytössä lähietäisyydeltä tapahtuvassa kaukokartoituksessa (Eugenio ym. 2020). Lyhytaaltoista infrapunasäteilyä (SWIR, short-wave infrared) tallentavia sensoreita käytetään harvemmin, sillä ne ovat usein VNIR (visible and near-infrared) -sensoreita painavampia, kalliimpia ja suorituskyvyltään epävarmempia (Jenal ym. 2019). Niiden kehitykselle on kuitenkin kysyntää metsien monitoroinnissa, sillä havu- ja lehtipuulajit ovat erotettavissa erityisesti SWIR-aallonpituuksilla. SWIR-hyperspektrisensoreita ei myöskään ole markkinoilla samoissa määrin kuin VNIR-sensoreita. FPI-teknologiaa sovellettiin ensimmäistä kertaa SWIR-aallonpituuksille Mannilan ym. (2014) tutkimuksessa. Sitten sitä on hyödynnetty puulajien luokittelussa ainakin Honkavaaran ym. (2016) tutkimuksessa turvetuotantoalueen geometrian ja kosteusolosuhteiden selvittämiseen.

Hyperspektrikuvia nimitetään kerroksisen rakenteensa ja spatiaalisen luonteensa vuoksi usein datakuutioiksi. Hyperspektrisensoreilla kerätystä 2D-aineistosta

voidaan kaupallisten ohjelmistojen fotogrammetristen algoritmien avulla muodostaa jopa korkean resoluution 3D-pintamalleja (Aasen ja Bolten 2018), jos spatiaalisen sijainnin lisäksi korkeudesta on hyötyä. Hyperspektridatan digitaalisuus mahdollistaa aineiston hyvän saavutettavuuden ja pitkälle automatisoidun käsittelyn, joka lisää tulosten objektiivisuutta sekä toistettavuutta (Hennessy ym. 2020). Hyperspektrikuvat ovat paitsi suuria tietoaaineistoja myös visualisointiin sopivaa dataa niiden muodostuessa kymmenistä – jopa sadoista – päällekkäisistä kuvakerroksista, jotka on kuvattu samasta paikaista mutta eri aallonpituuskaistoilla alle kymmenen nanometrin välein (Grah ja Geladi 2007).

RPAS-laitteet ovat lähietäisyydellä lennätettäviä miehittämättömiä ilma-aluksia, joihin kiinnitettyjen sensoreiden avulla saadaan kerättyä yksityiskohtaista kaukokartoitusaineistoa. Kauko-ohjaaminen voi tapahtua joko manuaalisesti reaaliajassa tai etukäteen määritetyn kuvausohjelman mukaan automaattisesti. Kauko-ohjattavista ilma-aluksista kerättyjen aineistojen käyttö on 2000-luvulla yleistynyt nopeasti esimerkiksi kasvillisuuden ja biodiversiteetin tutkimuksessa, sillä ne ovat huokeampia tuottaa kuin korkean resoluution satelliitti- ja ilmakuvaineistot (Colomina ja Molina 2014). Niiden avulla voidaan täydentää mittauksia alueista, jotka jäävät kenttätyössä ja satelliiteista kuvattuna kattamatta (Eugenio ym. 2020). Lisäksi RPAS-laitteista on hyötyä ajallisesti, sillä samaa paikkaa voidaan monitoroida usein esimerkiksi saman tai peräkkäisten vuodenaikojen aikana.

RPAS-laitteet ovat usein akkukäyttöisiä, joten niiden kanssa on helppo operoida maastossa. Lennätyslaitteilla voidaan myös tehdä hyperspektrikuvausta haastavissa sääolosuhteissa ja pilvipeitteen alapuolelta, mikä vähentää datan käsittelytarvetta (Honkavaara ym. 2013). Toisaalta laitteiden akut eivät ole kovin pitkäkestoisia, joten lennätyksen tulee tapahtua tiettyjen aikarajojen puitteissa tai laitetta joudutaan lataamaan maastossa. Kylmä ilma voi osaltaan vaikuttaa akkujen kestävyysajan ja laitteen toiminta-aikaan. Laitteiston kestävyyttä ajatellen myös tuuli, kosteusolosuhteet ja ukkonen on syytä ottaa huomioon lennätystä suunniteltaessa. RPAS-laitteen lennätykseen vaikuttavat lisäksi alueelliset ilmailusuositukset ja -säädökset. Laitteet ovat matalan lennätyskorkeuden vuoksi

alttiita ulkoisille häiriötekijöille, kuten fyysisille esteille. Maksimilennätyskorkeus RPAS-laitteille on tällä hetkellä Suomessa 120 metriä (Traficom 2022).

Kiinnostus RPAS-laitteita kohtaan on vauhdittanut myös hyperspektrisensoreiden kehitystä, sillä sen myötä datan tallennukseen ja käsittelyyn käytettävät ohjelmistot ovat kehittyneet. Uudenaikaiset hyperspektrikamerat ovat niin pieniä ja kevyitä, että ne pystytään integroimaan RPAS-laitteeseen ja akun kesto sallii vielä kohtuullisen lennätysajan. Lentoaikaan vaikuttaa, sensorin keveyden ohella, RPAS-laitteen tyypin ominaisuudet ja lentoaika voi vaihdella 15 minuutista pariin tuntiin (Xue & Su 2017). Hyperspektrikameroiden ostohinnat ja käyttökustannukset ovat alentuneet, joten niistä on tullut saavutettavampia pienemmille toimijoille (Nezami ym. 2020). Hyperspektrikamerat ovat kuitenkin edelleen kalliimpia ja kookkaampia kuin monet muut kaukokartoitussensorit. RPAS-laitteella voidaan yhden lennätysajan aikana kuvata vain suhteellisen pieni alue, joten esimerkiksi kokonaisten valtioiden tutkiminen näiden kahden yhdistelmällä on toistaiseksi työläästä tai miltei mahdotonta pienelle tutkimusryhmälle. Hyperspektrikuvaus RPAS-laitteesta soveltuu kuitenkin hyvin monenlaisiin paikallisiin ja alueellisiin tutkimustehtäviin ja menetelmälle löydetään jatkuvasti uusia sovelluskohteita. Kuvaus- ja lentolaitteiden jatkuva ja nopea teknologinen kehitys tekee menetelmästä potentiaalisen myös laajempien alueiden tutkimustehtäviin tulevaisuudessa.

2.2 Boreaalisten puiden spektrisiä ominaisuuksia

Metsän puulajikoostumus vaikuttaa puustosta mitattaviin spektreihin. Rautiainen ja Lukeš (2015) mittasivat kuusimetsikön heijastavuutta kasvukauden aikana ja havaitsivat vähäistä kasvua vihreän valon ja lähi-infrapun (NIR) aallonpituusalueilla. Koivuvoittoisessa metsässä heijastavuus NIR-alueella puolestaan kasvoi huomattavasti ja samalla punaisen valon heijastus väheni. Tämä osoittaa, että metsän puulajikoostumuksella on merkitystä sen spektrisiä ominaisuuksia tarkasteltaessa. Merkittävä osuus metsien albedosta johtuu lehtien

optisista ominaisuuksista (Lukeš ym. 2013), minkä vuoksi niiden tunteminen on olennaista metsien kaukokartoituksessa. Rautiaisen ym. (2018) mukaan boreaalisten metsien lehtipeitettä on tutkittu spektrometreillä sekä laboratorioissa että maastossa. Lehtipuiden optisista ominaisuuksista on kuitenkin olemassa enemmän tieteellistä tietoa kuin havupuiden, koska kooltaan pienten ja kapeiden havunneulasten mittaaminen on osoittautunut haastavaksi. Lehvästön optisten ominaisuuksien määrittämisessä on siis monia haasteita ja maastodatan kerääminen on yhä työlästä. Tutkimukset eivät useinkaan ole vertailukelpoisia, sillä niissä käytetyt aineistonhankinta- ja prosessointimenetelmät poikkeavat toisistaan (Rautiainen ym. 2018).

Havumetsien optiset ominaisuudet ovat ainutlaatuiset johtuen havupuiden neulasten monisironnasta ja latvusten rakenteesta (Rautiainen ja Stenberg 2005). Kuten muillakin kasveilla, havupuiden neulasten pigmentit saavat aikaan valon absorptiomaksimin näkyvän valon aallonpituuksilla. Neulasten reflektanssi kasvaa NIR- ja lyhyen aallonpituuden infrapunaa (SWIR-) aallonpituuksilla (Rautiainen ym. 2018). Vaikka havupuilla onkin todettu yhteisiä spektrisiä ominaisuuksia, osa havupuihin kuuluvista lajeista poikkeaa toisistaan absorptio-, transmittanssi- ja säteilyvoimakkuuksien perusteella (Lukeš ym. 2013, Hovi ym. 2017, Rautiainen ym. 2018). Esimerkiksi Lukešin ym. (2013) mukaan männyillä oli heidän tutkimuksessaan kuusia korkeampi reflektanssi NIR-aallonpituuksilla mitattaessa, vaikka lajien transmittanssit vastasivatkin toisiaan. Toisin sanoen kuusilla havaittiin siis mäntyjä korkeampi absorptanssi NIR-alueella.

Neulasista mitattuihin spektreihin vaikuttavat neulasten ikä sekä vuodenaikaisvaihtelut. Hovin ym. (2017) tekemässä tutkimuksessa saman vuoden neulaset käyttäytyivät optisesti samoin kuin lehtipuiden lehdet. Nuoria neulasia on kuitenkin vaikeaa mitata niiden hauraan rakenteen ja korkean vesipitoisuuden vuoksi (Rautiainen ym. 2018). Talvehtivien, vanhempien neulasten osalta fotosynteesin aktivoituminen keväällä kasvattaa neulasten klorofyllipitoisuutta, joten näkyvän valon aallonpituuksilla säteilyn (PAR, Photosynthetically active radiation) absorptio lisääntyy ja punaisen valon käänneaste muuttuu. Toisaalta

tämä tarkoittaa PRI-indeksin (Photochemical Reflectance Index) ja vihreän valon heijastumisen pienentymistä. Absorptio NIR-alueella lisääntyy kevästä syksyyn, minkä on epäilty johtuvan neulasten tärkkelyspitoisuuden kasvusta. Neulasten SWIR-alueen optisista ominaisuuksista on toistaiseksi vähäisesti tutkimustuloksia eri vuodenaikoina. Siirryttäessä kesästä alkusyksyyn heijastavuudessa, transmittanssissa tai absorptiossa ei tapahdu huomattavaa muutosta, mutta vähäiset muutokset ominaisuuksissa johtuvat fotosynteesin vähenemisestä ja ovat siten vastakkaisia kuin siirryttäessä kevästä kesään (Rautiainen ym. 2018).

Vuodenaikaisvaihtelut vaikuttavat tutkimusten mukaan myös lehtipuiden lehtien optisiin ominaisuuksiin. Hovi ym. (2017) havaitsivat vihreän ja punaisen valon absorption lisääntyvän aikaisin keväällä ja laskevan kasvukauden lopussa rauduskoivulla (*Betula pendula*) ja metsähaavalla (*Populus tremula*). Möttus ym. (2014) puolestaan tutkivat harmaalepän (*Alnus incana*) ja rauduskoivun spektrejä ja totesivat NIR-alueen absorption lisääntyvän voimakkaimmin keväällä ja maltillisesti läpi kasvukauden, mikä vastaa neulasten spektreistä tehtyjä havaintoja. Hovi ym. (2017) saivat saman tuloksen NIR- ja SWIR-alueiden mittauksista. NIR- ja SWIR-alueilla transmittanssi laski vastaavalla tavalla, joten lehtien heijastavuus pysyi lähes vakiona (Hovi ym. 2017). Möttus ym. (2014) havaitsivat lehtien klorofyllipitoisuudessa nousua kasvukauden lopulle saakka, vaikka aiemmissä tutkimuksissa kasvun oli havaittu taantuvan keskikesällä.

Puulajit ovat kaiken kaikkiaan spektrisiltä ominaisuuksiltaan melko samankaltaisia (Ruiliang 2009). Tämä on vaikeuttanut niiden luokittelua keskinkertaisen resoluution sensoreilla kerätyistä aineistoista. Nykyään tiedetään, että havu- ja lehtipuiden väliset spektriset erot tulevat parhaiten esiin SWIR-alueella. Neulaset heijastavat huonommin SWIR-säteilyä (Hovi ym. 2017, Rautiainen ym. 2018). Esimerkiksi Viinikan ym. (2020) mukaan metsähaapa (*Populus Tremula L.*) erottui selkeimmin männystä (*Pinus Sylvestris L.*), kuusesta (*Picea Abies L.*) ja koivusta (*Betula sp.*) punaisen reunan käännepisteen (724–727 nm) sekä lyhyen infrapunaa (1520–1564 nm ja 1684–1706 nm) aallonpituuksilla. Yleensä havupuiden neulasilla on lehtipuita alhaisempi transmittanssi koko spektrillä ja siten myös

kokonaisuudessaan korkeampi absorptio ja alhaisempi albedo (Lukeš ym. 2013, Rautiainen ym. 2018). On kuitenkin huomattava, että saman vuoden neulasten optiset ominaisuudet ovat lähempänä lehtipuiden lehdistä tehtyjä havaintoja.

Borealisissa metsissä puiden latvusto on sen verran harvaa, että maaperän heijastuksella on vaikutusta latvustosta mitattuun heijastavuuteen (Markiet ja Möttus 2020). Rautiaisen ja Lukešin (2015) tutkimuksen mukaan maaperä ja aluskasvillisuus vaikuttivat merkittävästi boreaalisen metsäalueen reflektanssiin läpi koko kasvukauden. Tutkijat totesivat vaikutuksen olevan suurin näkyvän valon aallonpituuksilla, ja vaikutus jopa kasvoi osalla aallonpituusalueista kasvukauden huippua lähestyttäessä – toisin kuin voisi kuvitella lehvästön tihenemisen vuoksi. Markiet ja Möttus (2020) totesivat, että boreaalisen metsämaaperän heijastavuus kasvoi suhteessa tutkimusalan hedelmällisyyteen lähi-infrapun aallonpituuksilla, mikä vastasi hyvin kenttämittauksista saatuja tuloksia.

2.3 Hyperspektrikuvien esikäsittely

Hyperspektrikuvien analysointi on vaativaa, sillä siihen vaikuttavat monet tekijät, kuten kuvaussensorin ominaisuudet, valon intensiteetti (mm. varjot) ja tulokulma, luokittelumenetelmän suoriutuminen sekä referenssiaineiston laatu (Dadon ym. 2019). Hyperspektrikuvien analysoinnissa on oleellista huomioida valaistusolosuhteiden vaihtelu tilassa ja ajassa (Ruiliang 2009). Pikselit, jotka sijaitsevat kuvien varjoisissa kohdissa, voivat sisältää kuvauskohteen, sen lähellä sijaitsevien objektien sekä ilmakehän siroamaa säteilyä, ja suoran auringonvalon puuttuessa näillä pikseleillä on matala signaali-kohinasuhde (Trier ym. 2018). Datakuutioiden esikäsittelymenetelmillä pystytään onneksi jo minimoimaan tällaisia valon sironnan ja valaistusolosuhteiden aiheuttamia virheitä melko hyvin. Hyperspektrikuvien käsittelyketju sisältääkin yleensä aineiston radiometrinen ominaisuuksien säätämisen tasalaatuisiksi, mikä parantaa luokittelun

toistettavuutta eri alueilla ja erilaisissa valaistusolosuhteissa, ja sen lisäksi piirteiden eristämisen, valinnan ja luokittelun (Nezami ym. 2020).

Hyperspektrikuville tulee tehdä radiometrinen korjaus. Radiometrisellä normalisoinnilla pyritään vähentämään esimerkiksi ilmakehän, valaistuksen ja sensorin kohinan vaikutusta maanpinnasta ja kasvillisuudesta mitattuihin intensiteettiarvoihin (Hong ja Zhang 2006). Esimerkiksi Tuominen ym. (2017) totesivat, että puun suku ja laji olivat paremmin luokiteltavissa radiometrisesti kalibroittujen reflektanssien kuin korjaamattomien numeeristen pikseliarvojen (DN, Digital Number) perusteella. Hyperspektrikuvien radiometrisellä prosessoinnilla pyritään säätämään kuvat mahdollisimman tasalaatuisiksi, jotta luokittelun toistettavuus paranisi erilaisissa kuvauspaikoissa ja valaistusolosuhteissa (Nezami ym. 2020). Hyperspektrikuvissa on myös geometrisiä mittausvirheitä, minkä vuoksi niistä tehdään usein ortomosaiikki, jossa monta kuvaa yhdistetään yhdeksi kuvaksi, jotta niiden geometrinen asteikko saadaan yhtenäiseksi.

Hyperspektrikuva-aineistolle voi olla tarpeellista tehdä korjaus, joka tasoittaa heijastusgeometriasta aiheutuvia virheitä. Aasen ja Boltenin (2018) mukaan sopiva heijastusjakaumafunktion (BRDF, Bidirectional Reflectance Distribution Function) malli valitaan kuvauskohteen mukaan ja sille lasketaan korjauskerroin (engl. multiplicative correction factor). Korjauksella voidaan vähentää havainnointisuunnasta riippuvaista heijastuksen intensiteetin vaihtelua. Tarkoituksena on saada reflektanssiarvot vastaamaan nadiirisuunnan, eli suoraan ylhäältä päin tulevan säteilyn, reflektanssikerrointa (Aasen ja Bolten 2018).

Taajuuskaistojen eristämällä ja valinnalla on merkitystä hyperspektrikuvien luokittelun onnistumisessa. Piirteiden valinnalla tarkoitetaan aineiston luokittelun kannalta olennaisimman informaation automaattista tai manuaalista valintaa heijastusmosaiikista, mihin vaikuttaa kyseessä oleva tutkimusongelma (Nezami ym. 2020). Nezamin ym. (2020) mukaan piirteiden valinnan ajatellaan usein tarkoittavan sopivan luokittelijan valintaa, mutta sen sijaan kysymyksessä on sopivien geometrinen tai spektristen tietojen erottaminen luokittelua varten. Tähän

voidaan käyttää esimerkiksi pääkomponenttianalyysia (PCA). Sen tarkoitus on yksinkertaistaa aineistoa korrelaatioon perustuvalla piirteiden valinnalla säilyttäen silti mahdollisimman suuren osan aineiston alkuperäisestä varianssista (Dadon ym. 2019). Hyperspektridatan ulottuvuuksien karsimisesta ei kuitenkaan ole kaikkien tutkimusten perusteella hyötyä lajiluokittelussa. Mozgeriksen ym. (2018) mukaan PCA-analyysin käyttö spektripiirteiden valinnassa heikensi tulosten tarkkuutta esimerkiksi MLP- ja RF-luokittelijoilla (Mozgeris ym. 2018). Myös Nevalaisen ym. (2017) tutkimuksessa spektripiirteiden valinnan havaittiin heikentävän hieman luokittelun kokonaistarkkuuksia samoilla algoritmeilla.

Kasvillisuusindeksejä voidaan käyttää ja on useissa tutkimuksissa (mm. Cao ym. 2018, Wu ja Zhang 2019) käytetty hyperspektrikuvien pikseleiden valoisuuserojen tasoittamiseen ja aineiston piirteistämiseen. Kasvillisuusindeksit ovat algoritmeja, joiden avulla on mahdollista tutkia kasvillisuuden piirteitä erilaisilla sensoreilla kerätyistä kaukokartoituskuvista. Niiden avulla voidaan tehdä transformaatiota kasvillisuudesta mitatuille alkuperäisille heijastavuusarvoille (Cao ym. 2018). Tällöin saadaan kasvilajien eroavaisuuksia ja näkymättömiä piirteitä esiin. Caon ym. (2018) mukaan kasvillisuusindeksit hyödyntävät erilaisia suhdelukuja, erotuksia ja lineaarisia yhdistelmiä tähän sekä valoisuuserojen ja taustaympäristön vaikutusten minimointiin. Niille ei kuitenkaan ole olemassa yhteistä matemaattista lauseketta, sillä indeksejä käytetään hyvin monenlaisilla sensoreilla ja lennätysalustoilla kerättyjen aineistojen analysointiin (Xue ja Su 2017). Xue ja Su (2017) selventävät, että algoritmit on kehitetty spesifisesti tiettyjä sovelluskohteita varten ja niiden laskennallisissa kaavoissa hyödynnetään niin näkyvän valon kuin sen ulkopuolisia aallonpituuksia, joiden avulla pyritään saamaan edustavia mittauksia kasvillisuudesta.

2.4 Hyperspektrikuvien luokittelu ohjatun koneoppimisen menetelmillä

Luokittelulla viitataan prosessiin, jossa ensin valitaan tehtävään soveltuva luokittelija, säädetään mallin hyperparametrit (oppimista ohjaavat parametrit

luokittelijasta riippuen) sopiviksi testaamalla erilaisia mahdollisuuksia suhteessa referenssiaineistoon, koulutetaan valittu algoritmi, tehdään sille validointi ja tutkitaan sen toistettavuutta (Nezami ym. 2020). Luokittelua voidaan tehdä ilman koulutusdataa ohjaamattoman oppimisen menetelmillä tai, kuten tämän tutkielman tapauksessa, ohjatun oppimisen menetelmillä. Ohjatun oppimisen luokittelumenetelmät ovat tilastollisen oppimisen menetelmiä, joissa hyödynnetään niin sanottua koulutusaineistoa (engl. training data), eli joukkoa tunnettuja arvoja tutkittaville objekteille, tuloksia ennustavan mallin (engl. learner) rakentamisessa (Hastie ym. 2009). Algoritmit käyttävät funktiota, joka sisältää parametrejä kuvaamaan spektrisen syötteen ja sitä vastaavan luokka-arvon (engl. input ja output) välistä suhdetta. Parametrien perusteella koulutettu malli siis luokittelee sille annetun datan. Algoritmien avulla mallinnetaan datan ehdollista jakautumista (engl. conditional distribution) ja tuloksena on luokittelukartta (Alcolea ym. 2020). Ohjatun oppimisen menetelmien haaste hyperspektrikuvien luokittelussa on niiden vaatima suuri määrä mittausaineistoa, jonka kerääminen ja analysointi on työlästä (Dadon ym. 2019). Ohjatun oppimisen menetelmillä päästään kuitenkin ohjaamattoman oppimisen menetelmiä tarkempiin tuloksiin luokiteltaessa hyperspektrikuvia, sillä ohjatun oppimisen malleille annetaan koulutusvaiheessa tietoja luokkakohtaisesti (Alcolea ym. 2020).

Puulajien luokittelussa voidaan käyttää lukuisia erilaisia algoritmeja, joista tähän tutkielmaan valittiin tarkasteltavaksi muutamia. RF (Random Forest) on yksi suosituimmista, jollei suosituin, puulajien tutkimuksessa käytetty luokittelija. RF rakentuu joukosta keskenään korreloimattomia luokittelupuita (engl. decision trees), joista jokainen äänestää ensin luokiteltavan näytteen luokkaa, minkä jälkeen malli päätyy ennustamaan näytteen eniten ääniä saaneeseen luokkaan (Breiman 2001, Raczko ja Zagajewski 2017, Alcolea ym. 2020). RF:n luokittelupuut opetetaan tunnistamaan näytteen luokka satunnaisten alkuperäisten spektrien ja koulutusaineiston havaintojen avulla. Siten luokittelupuista tulee keskenään erilaisia, mutta ne pystyvät koulutuksen ansiosta luokittelemaan alkuperäisestä aineistosta kuitenkin suurimman osan näytteistä oikeaan luokkaan.

LDA-luokittelija on yksinkertainen parametrinen malli, jonka toimintaperiaate perustuu luokkien keskipisteiden välisen etäisyyden maksimointiin ja toisaalta luokan sisäisten varianssien minimointiin (Cao ym. 2018). Menetelmä kuvaa luokkiin kuuluvien havaintojen joukkoa moniulotteisten normaalijakaumien avulla, joten tämän vuoksi LDA olettaa aineiston normaalijakautuneisuutta ja muuttujien homoskedastisuutta eli sitä, että toisen muuttujan varianssi pysyy suhteellisen vakiona suhteessa toiseen muuttujaan.

Myös SVM (Support Vector Machine) -algoritmi on hyvin suosittu menetelmä puulajien tunnistuksessa. SVM-luokittelijat pyrkivät löytämään optimaalisen piirreavaruuden (tai sen laajennoksen) sisäisen hypertason, jolla saavutaan korkein marginaali/etäisyys toisistaan eroteltavien luokkien välille (Raczko ja Zagajewski 2017). SVM on kernel-funktiopohjainen ja kaksiluokkainen menetelmä, jossa luokittelijan omaan luokkaan kuuluva aineisto saa positiivisen arvon ja loput aineistosta negatiivisen (Alcolea ym. 2020). Toisin sanoen näytteille tehdään ydinfunktiolla muutos, minkä jälkeen malli pyrkii erottamaan ne lineaarisesti kahteen luokkaan optimoiden luokkien välisen marginaalin (Melgani ja Bruzzone 2004). SVM-algoritmia voidaan kuitenkin käyttää myös useampi- kuin kaksiluokkaiselle aineistolle. SVM-luokittelijan suorituskyvyn on todettu olevan hyvä verrattuna muihin suosittuihin algoritmeihin sellaisillakin aineistoilla, joissa koulutusnäytteitä on vähäisesti ja luokat ovat heterogeenisiä (Melgani ja Bruzzone 2004). Melganin ja Bruzzoneen (2004) mukaan tämä perustuu luokittelijan toimintaperiaatteeseen, joka ei vaadi luokkien tilastollisen jakauman estimointia. He kuitenkin huomauttavat, että hyperspektrikuvia käsiteltäessä haasteita voi tulla näytteiden jakamisessa lukuisiin keskenään spektrisesti toisiaan muistuttaviin luokkiin, sillä algoritmi on alun perin suunniteltu kaksiluokkaiseen luokitteluun (engl. binary classification). Moniluokkainen SVM-luokittelu voi siten johtaa hieman alhaisempiin luokittelutarkkuuksiin kuin kaksiluokkainen menetelmä (Melgani ja Bruzzone 2004).

Yksi vaihtoehto kuvien käsittelyyn ovat suositut NN (Neural Network) -menetelmät, jotka ovat epälineaarisia ja kaksivaiheisia (engl. two-stage) regressio-

tai luokittelumalleja (Hastie ym. 2009). Ne perustuvat neuronien verkostoihin, joissa eri kerrokset (engl. layer) ovat linkittyneet sekvensoidusti toisiinsa (Alcolea ym. 2020). MLP (Multi Layer Perceptron) on NN-malli, jossa on vähintään yksi piilokerros *input*- ja *output*-tasojen välissä (Alcolea ym. 2020). Kiranyazin ym. (2021) mukaan mallissa piilotetun tason tekoneuronit (engl. artificial neurons) tekevät monimuuttujaiselle syöteaineistolle lineaarisen muunnoksen, jota seuraa vielä, mahdollisesti epälineaarinen muunnos, esimerkiksi kynnystyys (engl. thresholding). Toimintaperiaate on sama kuin biologisessa neuronissa, jossa synapsit vastaanottavat toisten neuronien signaalit, minkä jälkeen prosessoidut signaalit yhdistetään soomassa ja lopulta aksonissa eli viejähaarakassa tapahtuu aktivaatio, jos soomasta tulevassa signaalissa on riittävän korkea potentiaali (Kiranyaz ym. 2021). MLP-menetelmää ei ole käytetty puulajien luokittelussa yhtä laaja-alaisesti kuin esimerkiksi RF-luokittelua (Nevalainen ym. 2017). Tähän voi vaikuttaa se, että MLP-luokittelijalla saatavien tulosten tulkinta on usein haastavampaa kuin yksinkertaisella RF-mallilla. LDA-, MLP-, SVM- ja RF-luokittelijoiden yhteisinä etuina voidaan kuitenkin pitää niiden nopeutta ja suhteellisen helppoa koulutettavuutta luokittelussa.

Luokittelussa voidaan myös käyttää syväoppimiseen (engl. deep learning) perustuvia algoritmeja. Syväoppimismenetelmät perustuvat useilla tasoilla tapahtuvaan prosessointiin ikään kuin ihmisäivoissa tapahtuvassa neuroniverkostojen työskentelyssä (Chen ym. 2014). Sisään syötettävien näytteiden ja tulosten välissä on useita tasoja käsittävä neuronien verkosto, jonka monimutkaisuus riippuu syöteaineiston esikäsittely- ja piirteiden eristämistarpeista. NN-luokittelijoita käyttämällä voidaan rakentaa syväoppimisen malleja. CNN (Convolutional Neural Network) on NN-malli, jossa on yksi tai usempia konvoluutio- ja *pooling*-kerroksia (Fricker ym. 2019, Alcolea ym. 2020). Konvoluutiokerrokset tunnistavat syöteaineistosta piirteitä, ja *pooling*-kerroksen tehtävä on vähentää seuraavaan kerrokseen syötettävän datan määrää. CNN-mallin käyttö puulajien luokittelun tutkimisessa hyperspektrikuvista on yleistynyt lupaavien tulosten vuoksi. CNN-luokittelija voi hyödyntää spektrisen profiilin

lisäksi tietoa pikseleiden spatiaalisesta sijoittumisesta toisiinsa nähden, ja ne sopivatkin hyvin moniulotteisen tiedon käsittelyyn muiden syväoppimis-menetelmien tapaan (Fricker ym. 2019).

Kiranyazin ym. (2021) mukaan syväoppimisen 2D- ja 3D-CNN-menetelmät ovat kiinnostaneet tutkijoita korkeiden luokittelutarkkuuksien innoittamina mutta ne vaativat paljon koulutusaineistoa ja niiden kouluttaminen suuren laskentakapasiteetin. Tämä on tilanne etenkin tutkittaessa yksiulotteisia signaaleja, kun koulutusdataa on vähäisesti ja sovelluskohde spesifinen. Tällöin on kuitenkin mahdollista käyttää laskennallisesti yksinkertaisempaa 1D-CNN-menetelmää, jolla onnistuu esimerkiksi vain spektrin aallonpituuksien (λ) intensiteettejä sisältävän aineiston tutkimus. 1D-CNN-mallit koostuvat useista CNN- ja MLP-tasoista. CNN-tasoissa tapahtuu 1D-konvoluutioita, jotka ovat yksinkertaistettuna lineaarisesti painotettuja summia 1D-signaalin pohjalta muodostetun opetusaineiston lähellä toisiaan olevien aallonpituuksien arvoista. CNN-tasossa käytössä ovat myös aktivaatiofunktio ja *pooling*-funktio. 1D-mallissa käytetään yksiulotteisia matriiseja eli vektoreita niin konvoluution ydinfunktioissa kuin piirrekartoissa. CNN-kerros koulutetaan erottamaan 1D-raakadadatasta piirteet, joita käytetään MLP-luokittelussa. Tiheät MLP-tasot toimivat samaan tapaan kuin kyseisessä algoritmossa. 1D-CNN-algoritmi yhdistääkin piirteiden erottamisen ja luokittelun samaan prosessiin, joka pyritään mallin koulutusvaiheessa optimoimaan parhaan mahdollisen luokittelutarkkuuden saavuttamiseksi. Muutama vuosi sitten kehitetty 1D-CNN-menetelmä ei siten laskennallisesti ole kovin paljon monimutkaisempi kuin tutkimuksissa enemmän käytetyt luokittelijat, kuten RF, SVM ja MPL, mutta sillä voi olla mahdollista saavuttaa korkeampia luokittelutarkkuuksia (Kiranyaz ym. 2021).

Hyperspektrikuvien on todettu monissa tutkimuksissa antavan luotettavampia tuloksia ohjatun oppimisen luokitteluissa kuin muiden kaukokartoitusaineistojen. Mozgeriksen ym. (2018) mukaan verrattaessa hyperspektri- ja CIR-kuvia erot luokittelutarkkuuksissa olivat tilastollisesti merkitseviä ja osoittivat hyperspektrikuvien suoriutuvan paremmin. Toisaalta osa tutkimustuloksista osoittaa, ettei

hyperspektrikuvaus aina ole osuvin aineistonkeruumenetelmä luokittelua varten. Zhao ym. (2020) vertailivat latvuksen laserskannaukseen (LiDAR) ja pikselipohjaiseen hyperspektrikuvaukseen pohjautuvaa puulajiluokittelun tarkkuutta. Latvuksen laserskannaus osoittautui tarkemmaksi menetelmäksi niin lehtipuulajien kuin puulajiluokkien tunnistamisessa (Zhao ym. 2020).

Hyperspektrikuvien ja muiden kaukokartoitusaineistojen yhdistämisen on myös todettu useasti parantavan luokittelun tarkkuutta ohjatun koneoppimisen menetelmillä. Sothe ym. (2020) totesivat pistepilvistä saatujen 3D-piirteiden yhdistämisen hyperspektrikuviin lisäävän luokittelutarkkuutta ohjatun oppimisen menetelmillä jopa 13–17 prosenttia riippuen luokittelijasta (SVM tai RF) ja tutkimusalueen ominaisuuksista (Sothe ym. 2020). Mozgeris ym. (2018) totesivat infrapunakuvien yhdistämisen 16 spektrikanavan hyperspektriaineistoon kasvattavan luokittelutarkkuutta tarkimmalla luokittelijalla (MLP), vaikka tulos ei ollutkaan tilastollisesti merkitsevä. Vastaavaa vaikutusta ei havaittu 64 spektrikanavan kohdalla (Mozgeris ym. 2018). Nezamin ym. (2020) tutkimuksessa hyperspektri- ja RGB-kuvien sekä rakenteellisen aineiston yhdistämisellä yllettiin parhaisiin luokittelutarkkuuksiin CNN-luokittelijalla.

2.5 Tutkimustuloksia hyperspektrikuvien käytöstä puulajien luokittelussa

Hyperspektrikamerat tallentavat valon spektristä paljon kapeita aallonpituuskaistoja ja ovat siten spektrisesti hyvin tarkkoja. Niiden käyttöönnoton myötä voidaankin yhä tarkemmin tunnistaa aallonpituuksia, joilla puulajit eroavat toisistaan. Kun spektrinen informaatio lisääntyy, puulajiluokittelun tarkkuus paranee (Dadon ym. 2019). Sensoreiden ohella puulajien luokittelussa käytettävien algoritmien on oltava riittävän kehittyneitä huomioidakseen monia muuttujia, kuten puiden koon, muodon, väri-intensiteetin ja muut spektriset sekä geometriset ominaisuudet. Sensoreiden spatiaalisen tarkkuuden kehitys on puolestaan oleellista tarkan koulutusdatan keräämisessä koneoppimisen luokittelijoita varten (Fricker ym. 2019). Kuvien sisältämän aineiston muuttuminen

moniuolotteisemmaksi johtaa väistämättä myös niiden käsittelyn monimutkaistumiseen (Dadon ym. 2019).

Puulajiluokittelusta tehtyjen tutkimusten vertailtavuudessa on monia haasteita, joita aiheuttavat erilaiset aineiston keruu- ja analysointimenetelmät, tutkimuksen kohteena olevien puulajien määrät, erilaiset ominaisuudet ja tiheydet sekä hyperspektrikuvien toisistaan poikkeavat spektriset ja spatiaaliset skaalat (Viinikka ym. 2020). Hennessyn ym. (2020) tutkimus osoitti, että spektriipiirteiden valinnasta saatujen tulosten yleistettävyyteen tulisi myös suhtautua varauksella. Liiallisten yleistysten tekemisestä puulajiluokittelun tutkimusprosesseihin liittyen olisikin toistaiseksi hyvä pidättäytyä ja suhtautua jokaiseen tutkimukseen representaationa tietynlaisista lajeista, tutkimusalueista ja menetelmistä. Tutkimusta tällä saralla tarvitaan lisää, jotta tutkimusprosesseista saataisiin vertailukelpoisempia.

Clark (2020) vertaili tuoreessa tutkimuksessaan satelliiteista otettujen multi- ja hyperspektrikuvien tarkkuutta metsätyyppien luokittelussa. Hän totesi, että simuloidulla hyperspektrisatelliittidatalla oli 8,2–12,6 % korkeampi luokittelun kokonaistarkkuus kuin todellisilla Landsat 8- ja Sentinel-2-multispektrisatelliittikuvilla. Toisaalta simuloidulle Sentinel-2-aineistolle saatiin korkeimmaksi luokittelun kokonaistarkkuudeksi 74,3 % (Kappa -arvo 0,70) vuodenajasta riippumattomilla taajuuskaistoilla ja indekseillä, mikä oli vain 1,6 % matalampi kuin simuloiduilla hyperspektrikuvilla (Clark 2020). Tämä osoittaa, että satelliiteista otettujen multispektrikuvien avulla ylletään tyydyttäviin tarkkuuksiin metsätyyppien luokittelussa. Voidaan kuitenkin pitää todennäköisenä, että sensorin spektrisen ja spatiaalisen resoluution kasvattaminen parantaisi metsiköiden tai yksittäisten puiden luokittelun luotettavuutta satelliiteista käsin.

Saarinen ym. (2017) tutkivat fotogrammetristen pistepilvien ja hyperspektrikuvien soveltuvuutta biodiversiteetti-indikaattoreiden kartoittamiseen eteläisissä boreaalisissa metsissä. Käytetyt indikaattorit olivat rakenteellinen monimuotoisuus sekä kuolleiden puiden ja havupuiden määrä. Metsikkötason kuvamateriaali hankittiin UAV-laitteella, johon oli kiinnitetty sekä FPI-hyperspektrikamera että RGB-kamera. Analyysissä käytettiin *k*-NN-metodia tutkittavien puiden lajin

(mänty, kuusi, koivu, haapa ja leppä), rinnankorkeuden läpimitan, korkeuden ja elossaolon ennustamiseen. Mallinnuksessa käytettäviä spektripiirteitä valittaessa tutkijat käyttivät RF-luokittelijaa ja Pearsonin korrelaatiokerrointa. Kaikkien kanavien keskiarvospektri jokaiselle puulajille laskettiin elossa olevista puuyksilöistä. Yksittäisten puiden lajiluokittelun tarkkuuden keskiarvoksi saatiin 64,3 %. Tämä johtui ensinnäkin liian suurista tarkkuuksista paikoissa, joissa puun latvuksen rajaaminen ja pistepilven tunkeutuminen maanpinnan tasolle epäonnistui. Toisaalta tarkkuus sai liian pieniä arvoja paikoissa, joissa korkeimmat puut peittivät niiden alapuolella olevien puiden latvuksia. Kartoitustarkkuuksien suhteellinen keskivirhe oli pienempi paikoissa, joissa lehtipuita oli keskimääräisesti vähemmän (Saarinen ym. 2017). Tämä osoittaa, että metsän tiheydellä ja lajikoostumuksella on vaikutusta puulajiluokittelun onnistumiseen.

Ballantin ym. (2016) tutkimuksessa, jossa kuvattiin korkean biodiversiteetin metsäalueita, sekä SVM että RF ylsivät kaikissa kahdeksan puulajin luokittelussa yli 90 % tarkkuuksiin. He havaitsivat, että koulutusaineiston kasvattaminen paransi SVM-luokittelijan omaa suoriutumista ja suoriutumista suhteessa RF-luokittelijaan. SVM-luokittelijalla pystyttiin myös Viinikan ym. (2020) tutkimuksessa luokittelemaan tarkemmin haavat muista boreaalisen metsän puulajeista kuin RF-luokittelijalla (SVM:llä kokonaistarkkuus 84 % ja Kappa-arvo 0,74). Heidän tutkimuksessaan paras luokittelun kokonaistarkkuus (92 %) saavutettiin käyttämällä keskimääräisiä reflektanssiarvoja SVM-mallissa ja yhdistämällä siihen kasvillisuusindeksejä. Tutkimuksen perusteella voidaan kuitenkin todeta, että luokittelutarkkuudet haavalle, koivulle, männylle ja kuuselle olivat hyviä molemmilla luokittelijoilla: kokonaistarkkuudet 70–84 % ja Kappa-arvot 0,59–0,78 (Viinikka ym. 2020). Caon ym. (2020) vertaillessa kolmea eri koneoppimisen luokittelijaa kahdeksan eri mangrovepuun luokittelussa SVM-luokittelija ylsi kaikista korkeimpaan luokittelutarkkuuteen tuloksen ollessa jopa 93,13 %. Laskennallisten piirteiden, kuten kasvillisuusindeksien, käyttö kasvatti luokittelun tarkkuutta joitakin kymmenyksiä (Cao ym. 2020).

Nevalainen ym. (2017) kuvasivat FPI-hyperspektrikameralla ja RGB-kameralla kattavasti eri-ikäisiä ja eri lajeja edustavia puuyksilöitä borealisessa metsässä. Tutkimuksen tavoitteena oli selvittää 3D-pistepilvien ja erilaisten spektri- ja piirteiden suorituskäytännön yksittäisten puiden havaitsemisessa ja lajiluokittelussa. Tutkijat käyttivät C4.5- (decision trees), RF-, k-NN- (*k*-Nearest Neighbors, *k*-arvoilla 1, 3, 7), MLP- ja Naive Bayes -luokittelijoita koivun, kuusen, männyn ja lehtikuusen luokittelussa. Testissä parhaiten suoriutuivat MLP ja RF, 94,9 % ja 95,2 % kokonaistarkkuuksilla (Nevalainen ym. 2017). Mozgeris ym. (2018) puolestaan kuvasivat Bekas X32 ultrakevyeseen lentokoneeseen kiinnitetyillä Rikolan hyperspektrikameralla ja Nikonin D800E CIR (color-infrared) -kameralla yhtäaikaaisesti urbaanien viherympäristöjen lehtipuulajeja (valeakaasia, rauduskoivu, hevoscastanja, metsävaahtera, saarnivaahtera ja metsälehmus). He käyttivät samoja luokittelijoita kuin Nevalaisen ym. (2017) tutkimuksessa ja totesivat, että parhaaseen luokittelutarkkuuteen riippumatta puulajista tai käytettyjen spektrikanavien määrästä päästiin MLP-luokittelijalla, jonka lajikohtaiset luokittelun tarkkuudet vaihtelivat 51–72 % välillä (Mozgeris ym. 2018). Nevalaisen ym. (2017) tutkimuksessa käytettäessä MLP-luokittelijaa ja piirteiden valintaa saatiin heikoimpia tuloksia lehtikuusen luokittelussa. Kaikkein yleisin luokitteluvirhe tutkimuksessa oli kuitenkin kuusien luokittelu männnyiksi. Tämä saattoi johtua osin siitä, että opetusaineisto sisälsi enemmän mäntyjä kuin kuusia. Koivujen kohdalla toistettavuuden (engl. recall) ja tarkkuuden (engl. precision) arvot olivat yli 95 %, sillä niiden spektrit erosivat merkittävästi havupuiden spektreistä etenkin lähi-infrapun aallonpituuksilla (Nevalainen ym. 2017).

Huolimatta RF:n hyvästä suorituskäytännöstä puulajiluokittelussa on osassa tutkimuksista saavutettu parempia tuloksia vaihtoehtoisilla menetelmillä. Esimerkiksi Tuomisen ym. (2017) tutkimuksessa puulajien luokittelua tehtiin hyperspektri- sekä RGB-kuvien avulla. Aineisto koostui 26:sta Mustila Arboretum -tutkimusmetsässä kasvavasta puulajista. Aineisto kuvattiin kahdella FPI-hyperspektrikameralla aallonpituusväleillä 400–1000 nm (VNIR) sekä 1100–1600 nm (SWIR). RGB-kuva-aineistoa käytettiin tiheän 3D-läpisyavaruuden luomiseen. Data-analyysissä käytettiin k-NN-luokittelijaa ja geneettisiä algoritmeja (engl. genetic

algorithm, GA) yhtäaikaaisesti sekä vaihtoehtoista RF-metodia. Metodien vertailussa kävi ilmi, että k -NN:n ja GA:n yhdistelmä oli RF metodia tarkempi puulajin- ja suvun ennustamisessa. Ero oli 4–7 % lajin ja 2–5 % suvun mukaan luokiteltaessa. VNIR-aallonpituuksia käyttämällä päästiin parempaan estimointitarkkuuteen kuin SWIR-aallonpituuksilla. Paras luokittelun tarkkuus saavutettiin yhdistämällä VNIR-, SWIR- ja 3D-piirteitä. Tällöin oikein luokiteltujen puiden osuus oli 82,3 % lajin ja 86,9 % suvun mukaan (Tuominen ym. 2017).

Lisäksi Raczkon ja Zagajewskin (2017) verratessa kolmen eri luokittelijan soveltuvuutta puulajiluokitteluun hyperspektrikuvista, selvisi ANN-luokittelija (Artificial Neural Network) tehtävästä SVM- ja RF-luokittelijoita paremmin. Vertailussa olivat siis SVM-, RF- ja ANN-algoritmit, joiden avulla pyrittiin luokittelemaan tutkimusalueen viisi yleisintä puulajia: kuusi (*Picea Alba L. Karst*), lehtikuusi (*Larix decidua Mill.*), leppä (*Alnus Mill*), pyökki (*Fagus sylvatica L.*) ja koivu (*Betula pendula Roth*). Tutkijat havaitsivat, että ANN-luokittelijalla oli korkein kokonaistarkkuuden mediaani (77 %). Toisaalta sillä oli myös kahteen muuhun luokittelijaan verrattuna hieman korkeampi kokonaistarkkuuden varianssi (Raczko ja Zagajewski 2017).

Syväoppimismenetelmien hyödyntäminen yksittäisten puiden luokittelussa on myös viime vuosina nostanut päätään tutkimuksessa. Nezami ym. (2020) saivat 3D-CNN (3D-Convolutional Neural Networks) -menetelmällä parhaan luokittelun tuloksen koivulle (97,4 %), männylle (99,6 %) ja kuuselle (94,8 %) käyttäessään luokittelijaa, jossa oli yhdistetty hyperspektri- ja RGB-tasot. RGB-, hyperspektri- ja korkeusmallitasoilla sekä niiden eri yhdistelmillä 3D-CNN-menetelmän luokitteluvarmuus oli n. 5 % korkeampi kuin MLP:n (Nezami ym. 2020). Miyoshi ym. (2020) käyttivät syväoppimisen menetelmää, jossa he yhdistivät CNN-luokittelun ja parhaiten luokittelutehtävään soveltuvien taajuuskaistayhdistelmien etsimisen. Tuloksena oli 95,9 % luokittelutarkkuus, mikä oli huomattavasti korkeampi kuin käytettäessä PCA-menetelmää spektripiirteiden valinnassa (Miyoshi ym. 2020). Myös Trierin ym. (2018) tutkimuksessa syväoppimiseen perustuva NN-menetelmä suoriutui parhaiten (87 % tarkkuudella) koivun, männyn

ja kuusen luokittelussa. Fricker ym. (2019) saavuttivat CNN-luokittelijalla tarkkuuden keskiarvoksi 87 % luokitellessaan seitsemää eri puulajia, vaikka luokittelun tarkkuudessa olikin isoja lajikohtaisia eroja (67–95 %). Sothe ym. (2020) totesivat tutkimuksessaan, että luokiteltaessa puulajeja yksinomaan hyperspektritaajuuskaistojen perusteella CNN oli 22–26 prosenttia tarkempi kuin SVM ja RF. Tyypilliset koneoppimisen menetit osoittautuivat nopeammiksi kuin CNN mutta niiden kokonaistarkkuudet vaihtelivat enemmän eri lähdeaineistoilla, koska piirteytys ei tapahtunut automaattisesti (Sothe ym. 2020).

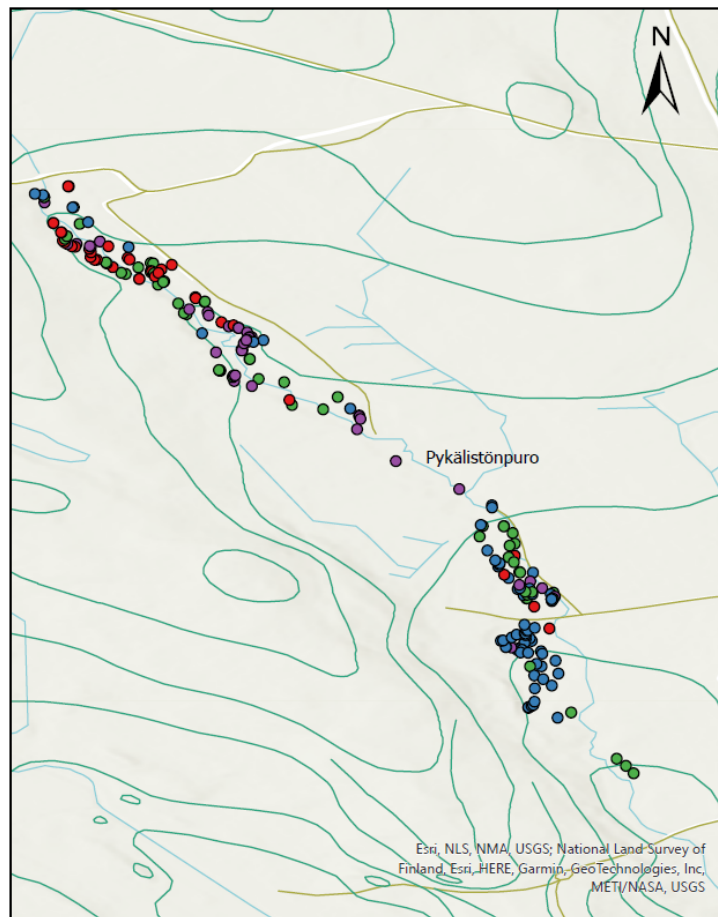
3 AINEISTO JA MENETELMÄT

3.1 Aineiston keruu, tutkimusalue ja välineet

Tätä tutkielmaa varten yritettiin kerätä omaa aineistoa Taka-Keljossa sijaitsevasta vanhasta boreaalisesta metsästä heinäkuussa ja syyskuussa vuosina 2018–2019 tapahtuneilla RPAS-lennoilla. Kuvausolosuhteet eivät kuitenkaan olleet lennätyspäivinä optimaaliset, vaikka tuulen voimakkuus ei sääennusteiden mukaan ollut yli 5 m/s. Puuskissa tuuli oli kuitenkin selvästi voimakkaampaa, ja siten kuvissa havaittiin visuaalisessa tarkastelussa merkittävää epätarkkuutta johtuen puiden oksien heilumisesta tuulella. Kun käytetyn hyperspektrikameran 55 taajuuskaistaa sisältävän kuvan ottoon meni aikaa yli yksi sekunti, olivat yksittäiset oksat eri taajuuskaistojen kuvissa eri asennoissa. Täten kuvien pikseleiden kohdistaminen ei onnistunut edes tyydyttävästi. Pikseleiden kohdistamiseen yritettiin kehitellä menetelmää, jossa sitä olisi tehty puuyksilökohtaisesti. ArcMap -ohjelmalla rajattiin puiden kruunuista monikulmioita, jotka oli tarkoitus poimia tiff-muotoisista hyperspektrikuvista R-koodin avulla, joka sitten jatkaisi kyseisten vain yhteen puuyksilöön kuuluvien pikseleiden kohdistusta. Tämä menetelmä havaittiin kuitenkin liian hitaaksi R-ohjelmalla toteutettavaksi eikä kameraan liittyvän Coregistering-ohjelman (v2.0, Senop Oy, Kangasala, Suomi) avullakaan puuyksilöiden pikseleiden

kohdistaminen onnistunut kunnolla. Täten itse otettujen kuvien esikäsittely todettiin liian paljon aikaa vieväksi ja monimutkaiseksi tehtäväksi tutkielman toteutusaikataulua ajatellen.

Edellä mainittujen haasteiden vuoksi tutkielman aineistona käytettiin lopulta toista Jyväskylän yliopiston hanketta varten kerättyä hyperspektrikuva-aineistoa. Aineisto oli kerätty Äänekosken Pykälistössä sijaitsevalta borealiselta metsäalueelta. Tutkimusalue sijaitsee Pykälistönpuron varressa (Kuva 1). Puron varsilla on usein edustettuna paitsi borealisille metsille tyypillisiä havupuulajeja myös erilaisia lehtipuita. Tässäkin tapauksessa tutkimusalueella tavataan lehtipuulajeja, kuten koivu, haapa, harmaaleppä, tervaleppä sekä raita.



Kuva 1. Tutkimusalueen kartta, johon merkatut pisteet ovat maastodataan merkattuja puita (punainen = mänty, sininen = kuusi, vihreä = koivu, violetti = harmaaleppä). Taustakartta-aineistot: Esri, NLS, NMA, USGS; National Land Survey of Finland, Esri, Garmin, Geo Technologies, Inc, METI/NASA ja USGS.

Tutkimusalueelta oli olemassa hyperspektrikuva-aineiston lisäksi maastoaineisto, jossa tämän tutkielman kannalta tärkeimpinä muuttujina oli merkattu tutkimuspuiden lajit ja niiden GPS-laitteella mitatut sijaintitiedot. Nämä on esitetty Kuvassa 1.

Hyperspektrikuvauksessa käytettiin Senop Oy:n (ent. Rikola) valmistamaa FPI-hyperspektrikameraa (Kuva 2). FPI-teknologia perustuu kahden vastakkaisen peilin väliin jäävän ”ilma-aukon” läpi päästämiin aallonpituuksiin, jotka vaihtelevat peilien välisen etäisyyden mukaan (Näsi ym. 2016). Tutkimuksessa käytetty kamera tallentaa 2D-kuvia näkyvän valon ja lähi-infrapunan aallonpituuksilla (VNIR; 400–1000 nm). Kuvauksen aallonpituusalue voi olla enintään 400 nm:n levyinen, ja jos kyseistä säätöaluetta haluttaisiin muuttaa, tulisi sen tapahtua valmistajan tehtaalla. Kuvauksissa tämä aallonpituusväli oli 500–900 nm. Äänekosken aineiston keruu oli tapahtunut 11.6.2020 lennättämällä kameraa ennalta määriteltyjä lentolinjoja pitkin tarkoituksena kuvata suorakulmion muotoinen alue Pykälistönpuron ympäriltä.



Kuva 2. Senop Oy:n FPI-hyperspektrikamera kiinnitettynä kolmijalkaan.

Hyperspektriraakadatan tallentamista varten laadittiin sekvenssi Rikola Hyperspectral Imager Software 2.0 -ohjelmalla (Senop Oy, Kangasala, Suomi). Sekvenssin avulla määritetään tallennettavien aallonpituuksien ja puoliarvoveveyden (FWHM, Full Width at Half Maximum) yhdistelmät ja niiden tallennusjärjestys. Puoliarvoveveys kertoo kameran spektrisestä erotuskyvystä. Tutkimuksessa käytettyjen kuvien puolileveysarvojen minimi oli 5,32 ja maksimi 13,88 nm. Kuvien resoluutio oli 1010 x 1010 pikseliä.



Kuva 3. DJI Matrice 600 Pro -kopteri, jonka kyydissä on hyperspektrikamera sekä säteilyvoimakkuus- ja GPS-mittarit. Kuva otettiin ensimmäisessä lennätyspaikassa Taka-Keljossa Jyväskylässä.

Hyperspektrikameran lennättäminen tapahtui kauko-ohjattavalla DJI Matrice 600 Pro -heksakopterilla (Kuva 3), jossa on kuusi roottoria. Se pystyy kantamaan 6 kg painolastin ja kopterin lentoaika ilman painolastia standardiakkuineen on 32 minuuttia. Tutkimuksessa käytetty FPI-kamera on kevyt (720 g), joten sen kiinnittäminen lentoalustaan ei lyhentänyt maksimilentoaikaa merkittävästi. Kopterissa oli kiinni Ronin-MX-gimbaali, joka tehtävä oli pitää sensori mahdollisimman vakaana lennätysten ajan, jotta kuvakuution eri taajuuskaistat pysyisivät linjassa toisiinsa nähden. Äänekoskella kopterin lentokorkeus oli 70 metriä lennätysten lähtöpaikkaan nähden, mutta korkeus maanpintaan nähden ei ollut lennätysten aikana koko ajan vakio. Maanpinnan korkeus laski ja siten etäisyys kopteriin kasvoi puronvartta alaspäin mentäessä. Kuvausolosuhteet olivat olleet Äänekosken lennätyspäivänä suhteellisen stabiilit valaistuksen suhteen, ja tuulta ei ollut havaittavissa lainkaan.

3.2 Aineisto

Kerätty aineisto koostuu yhteensä 97 hyperspektrispektrikuvasta. Tähän tutkimukseen soveltuvat kuvat valittiin referenssiaineiston ja kuvien laadun (pikseleiden valotus ja kohdistamisen onnistuminen) perusteella. Tutkimukseen valittiin lisäksi mukaan vain sellaisia kuvia, joissa olevien puiden laji pystyttiin todentamaan maastodatan perusteella. Tällaisia kuvia löydettiin 24 kappaletta. Tässä tutkimuksessa keskityttiin neljän puulajin luokitteluun, joista puolet oli lehtipuu- ja puolet havupuulajeja. Valitut puulajit olivat mänty, kuusi, koivu sekä harmaaleppä. Tutkimukseen valittujen puuyksilöiden lukumäärät ja pikseleiden keskimääräiset lukumäärät puulajeittain on lueteltu Taulukossa 1.

Taulukko 1. Tutkimuksessa käytettyjen puuyksilöiden ja pikseleiden keskimääräiset lukumäärät puulajeittain. Puuyksilöön kuuluvien pikseleiden lukumäärän vaihteluväli oli 328–6445 pikseliä kuusten ollessa tyypillisesti kapealatvaisempia.

Puulaji	Näytteiden lukumäärä	Pikseleiden keskimääräinen lukumäärä
Mänty	9	2450
Kuusi	9	1703
Koivu	13	2315
Harmaaleppä	10	2435

Puulajikohtaisia yksilömääriä tarkastelemalla Taulukosta 1 voidaan todeta, että kunkin puulajin edustajia oli mukana tutkimuksessa lähes yhtä paljon. Koivuista oli kuitenkin näytteitä vajaa kolmasosa enemmän kuin havupuulajeista ja noin neljäsosa enemmän kuin harmaalepistä. Voidaan kuitenkin olettaa, ettei eri

luokkien näytteiden määrällä ole vaikutusta havaittuihin tuloksiin kunkin puulajin luokittelutarkkuuden osalta.

3.3 Aineiston esikäsittelymenetelmät

Hyperspektrikuvauksesta saatu raakadata kalibroitiin Rikolan Hyperspectral Imager Software 2.0 -ohjelman avulla. Tällä ohjelmalla hyperspektrikuville tehtiin kalibrointi käyttäen pimeäkuva (engl. dark reference), joka oli otettu juuri ennen lennätystä. Sensorin lämpenemisestä aiheutuva pimeävirta aiheuttaa sen, että CCD-kennon detektoreihin muodostuu enemmän elektroneja kuin kennolle saapuvien fotonien takia tulisi, mikä aiheuttaa otettuihin kuviin pikseleittäin erilaiset määrät kohinaa (Yang ja Kan 2020). Pimeävirran vaikutusta otettuihin kuviin pyritään ehkäisemään vähentämällä hyperspektrikuvista pimeäkuva. Pimeäkuvan otossa käytettiin samoja asetuksia valotusajan osalta kuin hyperspektrikuvissa. Pimeäkuva otettiin peittämällä kameran linssi.

Honkavaaran ym. (2013) mukaan ideaalitulanteessa kaikki säteet kulkisivat hyperspektrikameran FPI-yksikön läpi samansuuntaisesti. Jotta kameran polttovälin ja tuloaukon (engl. entrance pupil) läpimitan suhdetta saataisiin parannettua, osa säteistä kulkee kuitenkin yksikön läpi enintään 10 asteen kulmassa (Honkavaara ym. 2013). Tämä aiheuttaa sen, että keskiaallonpituus pienenee kuvan reunoja kohti mentäessä. Tätä ilmiötä nimitetään aiheeseen liittyvässä kirjallisuudessa spektriseksi hymyksi (engl. spectral smile). Tässä tutkimuksessa spektrisen hymyn korjaus tehtiin valitsemalla mukaan ainoastaan puita, jotka olivat hyperspektrikuvan keskiosassa. Honkavaaran ym. (2013) mukaan kuvien pitkittäis- ja sivuttaispeiton ollessa vähintään 60 % ja puoliarvoleveyden 10–40 nm on kuvan nadiirikohdassa spektrisen hymyn vaikutus alle viisi nanometriä, mikä voidaan yleensä jättää huomiotta.

Aallonpituuksien sekvenssiin perustuvan hyperspektrikuvan (engl. frame-format hyperspectral image) taajuuskaistoilla on erilaiset sijainnit lentoalustan etenemisen lentoradalla tai esimerkiksi tuulesta johtuvan heilumisen vuoksi (Vakalopoulou ja

Karantzas 2014), joten niiden pikseleiden spatiaalisten erojen korjaaminen on tarpeellista silloinkin, kun drone pysäytetään kuvien oton ajaksi. Sekvenssiin perustuvassa menetelmässä taajuuskaistojen valotusajat eivät ole täysin samanaikaiset ja täten niiden välillä on hienoisia spatiaalisia eroja, joilla on merkitystä analyysin kannalta. Pikseleiden kohdistaminen tehtiin kaikkia tutkimuskuvien taajuuskaistoja käyttäen Coregistering 2.0 -ohjelmalla (Senop Oy, Kangasala, Suomi), joka laskee pikseleille samanlaiset sijainnit eri taajuuksilla peräkkäin kuvatuissa tasoissa. Mozgeriksen ym. (2018) mukaan Coregistering v1.1 -ohjelma toimii hyvin pikseleiden kohdistamisessa, jos lentokorkeus on ollut riittävä ja lentonopeus maltillinen. Tässä tutkimuksessa käytettiin vastaavan ohjelman uudempaa versiota, ja drone pysähtyi kuvanoton ajaksi. Ohjelmalla muunnettiin samalla kuvat tiff-muotoisiksi, jotta ne pystyttiin lukemaan sisään ArcGIS Pro -ohjelmaan georeferointia varten.

Hyperspektrikuvien georeferointi, eli niiden esittäminen spatiaalisessa sijainnissaan maanpinnalla, tehtiin manuaalisesti ArcGIS Pro 2.7.1 -ohjelman (ESRI Inc., Redlands, CA, USA) avulla. Tämä vaihe oli tärkeä, jotta maastodataan merkityt puut pystyttiin tunnistamaan hyperspektrikuvista. Hyperspektrikuvien keskipisteille oli olemassa sensorin tallentamat koordinaatit, joiden avulla kuvat visualisoitiin karttatasossa. Puut esitettiin karttatasossa maastodatan GPS-koordinaattien avulla. Maanmittauslaitoksen sivuilta ladattua ortokuva-aineistoa käyttäen hyperspektrikuvista etsittiin maastosta tunnistettavissa olevia kohteita. Kontrollipisteet luotiin valitsemalla piste ensin lähdetasosta (hyperspektrikuva) ja arvioimalla kyseisen pisteen sijainti referenssiaineistossa (ortokuva-aineisto). Georeferoinnissa suositetaan yleensä automatisoituja menetelmiä, toisin sanoen esimerkiksi kontrollipisteiden automaattista laskemista algoritmeilla ominaisspektriin (engl. spectral signature) perustuen, mutta ArcGIS Pro -ohjelma edellyttää, että tällöin lähde- ja referenssiaineiston tulee olla yhteneväisiä spatiaalisesti, ajallisesti sekä vuoden ajan, kuvan orientaation, kuvan skaalan ja taajuuskaistojen yhdistelmien suhteen. Tällöin hyperspektrikuvien ja referenssiaineiston keruun olisi ollut hyvä tapahtua samanaikaisesti ja niiden olisi tullut sisältää osin samanlaiset taajuuskaistat. Kuvien visuaalisen tulkinnan etu

automatisoituihin menetelmiin verrattuna on, että ihmisäivot pystyvät erottelmaan tehokkaasti kuvista erilaisia piirteitä, esimerkiksi valoja ja varjoja (Horning ym. 2010).

Georeferoinnin jälkeen kuvista valittiin tunnistettavissa olevat puut. Valinnassa kriteereinä olivat puun sijainti mahdollisimman lähellä kuvan keskustaa, puun elossa olo ja pikseleiden kohdistamisen onnistuminen kuvissa esiintyvistä lievista värivirheistä huolimatta. Valituista puuyksilöistä poimittiin kruunun valaistun osan keskipisteen koordinaatit ja sen säde lähimpään ulkoreunaan mitattiin ArcGIS Pro -ohjelmalla. Kruunun säde oli minimissään 10 pikselin mittainen.

Aineistolle tehtiin esikäsittelyä R-ohjelman avulla. Ensin CSV-muotoiseen Excel-taulukon tiedot luettiin ohjelmaan muuttujina hyperspektrikuvan nimi ja kuvasta löytyvän puuyksilön laji, sekä edellä mainitut säteen pituus pikseleinä sekä sen latvuksen valaistun osan keskikohdan x- ja y-koordinaatit. Lisäksi puulajin suhteen epävarmoina pidetyt puut saivat tunnisteon 1 ja muut tunnisteon 0, jotta puuyksilöiden tunnistuksen onnistumista kyettiin myöhemmin arvioimaan puiden spektrien avulla. Ohjaajani Anssi Lensun rakentama R-koodi poimi taulukossa olevien muuttujien perusteella kunkin puun määritellyt pikselit ja piirsi ne näkyviin eri väreillä kuin alkuperäisessä kuvassa, jotta puuyksilöiden pikseleiden poiminnan onnistuminen voitiin varmistaa. Koodin tarkoituksena oli seuraavaksi tuottaa puuyksilökohtaisten pikseliarvojen taajuuskaistakohtaisiin keskiarvoihin perustuva opetusdata, joka sisältäisi kunkin puun tunnistetiedot sekä sen pikseleistä lasketut aallonpituuskaistojen heijastavuusprofiilit. Tämän lisäksi opetusdatassa kokeiltiin käyttää peräkkäisten aallonpituuskaistojen heijastavuuksien erotuksia. Toisin sanoen aallonpituuskaistan heijastavuusarvosta vähennettiin aina sitä edellisen aallonpituuskaistan vastaava arvo, jotta saatiin tietää, onko aallonpituuskaistojen välillä heijastavuuksissa tapahtunut muutos kasvava vai laskeva. Tätä aallonpituuskaistojen heijastavuuksien erotusta käyttämällä saatiin tämän aineiston tapauksessa käytännössä samantyyppisiä arvoja kuin saataisiin käyttämällä aallonpituuskaistojen heijastavuuksien derivaattoja. Esimerkiksi Caon ym. (2018) tutkimuksessa derivaattoja käytettiin

onnistuneesti tuomaan esiin spektrien aaltomuodoissa tapahtuvia muutoksia ja huippuarvoja, jotta puulajeja pystyttiin erottamaan entistäkin tarkemmin toisistaan.

Aineiston esikäsittelyyn oli erilaisia vaihtoehtoja, jotka nimettiin itse nimikkeillä Drop.dark, Adjust, Adj.species ja Adj.center. Drop.dark-esikäsittely hylkäsi kunkin puun kaikkein tummimmista pikseleistä tietyn prosentuaalisen osan. Ajatuksena tässä oli, että nämä tummimmat pikselit saattoivat useissa tapauksissa edustaa jotakin muuta kuin puun lehvästää, kuten puun oksia tai maastoa puun alla. Luokittelun ulkopuolelle jätettävää pikseleiden määrää tai osuutta pystyttiin muuttamaan muuttamalla esikäsittelytavan poimintarajana käytetyn kvantiilin arvoa. Tässä tutkimuksessa puuyksilön pikseleistä hylättiin 25 % eli käytettiin kvantiilia 0,25. Adjust-esikäsittelyllä puuyksilöstä poimittujen pikseleiden (heijastuksen aallonpituusarvojen keskiarvoon perustuvat) kirkkausarvot pyrittiin tasoittamaan säätämällä arvot kaikissa pikseleissä samanlaisiksi kuin kunkin puun kirkkaimmalla pikselillä, jotta saman puun lehvästön valossa ja varjossa olevat lehdet olisivat kirkkaudeltaan likimain vastaavia, eivätkä valaistuserot tuottaisi aineistoon ylimääräistä vaihtelua.

Esikäsittelyssä kokeiltiin myös käyttää PCA-analyysiä aineiston ulottuvuuksien kaventamiseen. PCA-analyysin todettiin kuitenkin heikentävän luokittelutuloksia, mikä voi johtua käytössä olleesta aineistosta. Tutkimusaineistossa saattoi olla liian vähän näytteitä suhteessa muuttujien korkeaan määrään. PCA-analyysin heikkoon suoriutumiseen voi kuitenkin olla lukuisia muitakin syitä. Valaistuseroista kertovaa, ensimmäistä pääkomponenttia ei kokeiltu jättää pois, jolloin luokittelussa olisi käytetty pelkästään sävyeroista kertovia komponentteja. Komponenttien määrää ei myöskään kokeiltu optimoida vaan niistä kymmenen suurinta otettiin mukaan luokitteluun. Lisäksi PCA-analyysiä olisi voitu testata alkuperäisellä pikseliaineistolla. Tässä tutkimuksessa käytettiin yksinomaan puuyksilökohtaisia pikseleiden keskiarvoja. On mahdollista, että näillä muutoksilla esikäsittelytapa olisi toiminut paremmin.

Esikäsittelyssä testattiin myös kasvillisuusindeksien käyttöä. Tutkimukseen valittiin neljä indeksiä, jotka olivat Wun ja Zhangin (2019) tutkimuksessa mukana

parhaita kokonaistarkkuuksia tuottaneissa piirrejoukoissa käytettäessä SVM-luokittelijaa. Nämä indeksit olivat NDVI (Normalized Difference Vegetation Index), PRI (Photochemical Reflectance Index), GNDVI (Green Normalized Difference Vegetation Index) sekä PSRI (Plant Senescence Reflectance Index) ja niiden laskukaavat löytyvät Taulukosta 2, jossa esitetään kaavoissa tämän tutkimuksen kuva-aineiston alkuperäistä laskentakaavaa lähimmät aallonpituudet. Indeksit liittyvät kasvillisuuden klorofylli- ja karotenoidipitoisuuksiin ja kolme niistä hyödyntää myös lähi-infrapunan aallonpituutta (Wu ja Zhang 2019).

Taulukko 2. Tutkimuksessa käytetyt kasvillisuusindeksit ja niiden laskukaavat.

Kasvillisuusindeksi	Kaava
Normalized Difference Vegetation Index	$NDVI = \frac{(\rho_{512,96} - \rho_{530,18})}{(\rho_{512,96} + \rho_{530,18})}$
Photochemical Reflectance Index	$PRI = \frac{(\rho_{533,53} - \rho_{568,69})}{(\rho_{533,53} + \rho_{568,69})}$
Green Normalized Difference Vegetation Index	$GNDVI = \frac{(\rho_{799,56} - \rho_{547,76})}{(\rho_{799,56} + \rho_{547,76})}$
Plant Senescence Reflectance Index	$PSRI = \frac{(\rho_{681,46} - \rho_{505,72})}{\rho_{750,57}}$

Tutkimuksessa löydettiin aineistolle tehokas esikäsittelymenetelmä, jolla saatiin merkittävästi nostettua luokittelun kokonaistarkkuutta kaikilla luokittelijoilla. Adj.species-esikäsittelyllä saatiin tasoitettua kirkkauseroja saman lajin puuyksilöistä poimittujen pikseleiden välillä. Kyseinen toiminto skaalasi kunkin puulajin pikselit kyseisen lajin spektrien heijastavuusarvojen maksimiarvon perusteella. Tutkimuksessamme luokittelua tehtiin puille, joiden lajit olivat etukäteen tiedossa. Tämä esikäsittelymenetelmä vaatiikin, että luokiteltavien puiden lajit tunnetaan, mikä tekee siitä soveltuvan vain vastaavalla tavalla toteutettuihin tutkimuksiin. Tunteamattomalla aineistolla lajikohtaisia heijastavuuseroja voisi olla mahdollista tasoittaa joko asettamalla kunkin lajin kaikkien puuyksilöiden keskiarvon nolnaan tai tekemällä tasoitus kuvakohtaisesti

johonkin kuvasta löytyvään maksimiin tai poimimalla puuyksilöt kaikkien kuvien mosaiikista, jossa alkuperäisten kuvien valaistuserot olisivat jo minimoitu. Puuyksilökohtaista heijastavuuksien keskiarvon sijoittamista nolnaan kokeiltiin Adj.center-esikäsitteilyn avulla, joka laskee annetun raja-arvon yläpuolella oleville heijastavuusarvoille keskipisteen. Kuvakohtainen esikäsitteily olisi ollut hankalampaa toteuttaa kunnolla, sillä kaikista kuvista ei esimerkiksi löytynyt saman puulajin edustajia kirkkaassa auringonvalossa.

3.4 Puulajien luokittelumenetelmät

Luokittelussa käytettiin puukohtaisia piirrevektoreita, jotka olivat pikseleiden aallonpituuskaistojen heijastavuuksien keskiarvoja. Lisäksi kokeiltiin käyttää pelkästään spektrien peräkkäisten aallonpituuksien heijastavuuksien erotuksia sekä molempia edellä mainittuja samanaikaisesti. Kolmantena vaihtoehtona mukaan otettiin vielä tutkimukseen valitut kasvillisuusindeksit (Taulukko 2).

Puulajien luokittelussa käytetyt algoritmit olivat lineaarinen diskriminanttianalyysi (LDA), RF ja SVM (ks. aliluku 2.4). Tarkemmin vertailuvaiheessa käytettiin askeltavaa LDA-algoritmia, joka käyttää luokittelussa osajoukkoa alkuperäisistä muuttujista. Jokaisella askeleella mallin rakennuksessa arvioidaan, mitkä muuttujat kaikkien muuttujien joukosta toimivat parhaiten luokkien erottamisessa. Luokittelua tehtiin R:n caret-paketilla (Kuhn M. 2021. caret: Classification and Regression Training. R package version 6.0-90. <https://CRAN.R-project.org/package=caret>), johon kaikki tutkimuksessa käytetyt algoritmit sisältyivät. Caret-paketin etuna on, että sen avulla luokittelijoiden hyperparametreja voidaan optimoida ja sen tuloksena saadaan arvio luokittelijan suorituskyvystä, koska se tekee valituilla luokittelijoilla automaattisesti ristiinvalidointia eli käyttää määrättyä osaa aineistosta koulutuksessa ja jäljelle jäävää osaa luokittelun hyvyden estimointiin. Tässä tutkimuksessa käytettiin kymmenkertaista ristiinvalidointia (engl. ten-fold cross-validation), jossa algoritmi käyttää noin 90 % prosenttia aineistosta koulutukseen ja loput 10 % siitä

testaukseen. Luokittelun hyvyyden testaukseen käytettiin siis eri osaa aineistosta kuin koulutukseen jokaisen 10 opetus- ja testauskierroksen yhdistelmän aikana.

Caret-paketissa RF:n säädettävät hyperparametrit olivat puiden määrä (ntree) sekä satunnaisten muuttujien, toisin sanoen ennustajien, määrä (mtry), jota käytetään jokaisessa puussa päätöksen tekoon. Luokittelussa käytettiin caret-paketissa oletuksena olevia hyperparametrien arvojen skaaloja. Oletusarvona mtry-hyperparametrille on kaikkien muuttujien neliöjuuri ja maksimiarvona kaikkien muuttujien määrä. Puita mallissa oli 300 kappaletta.

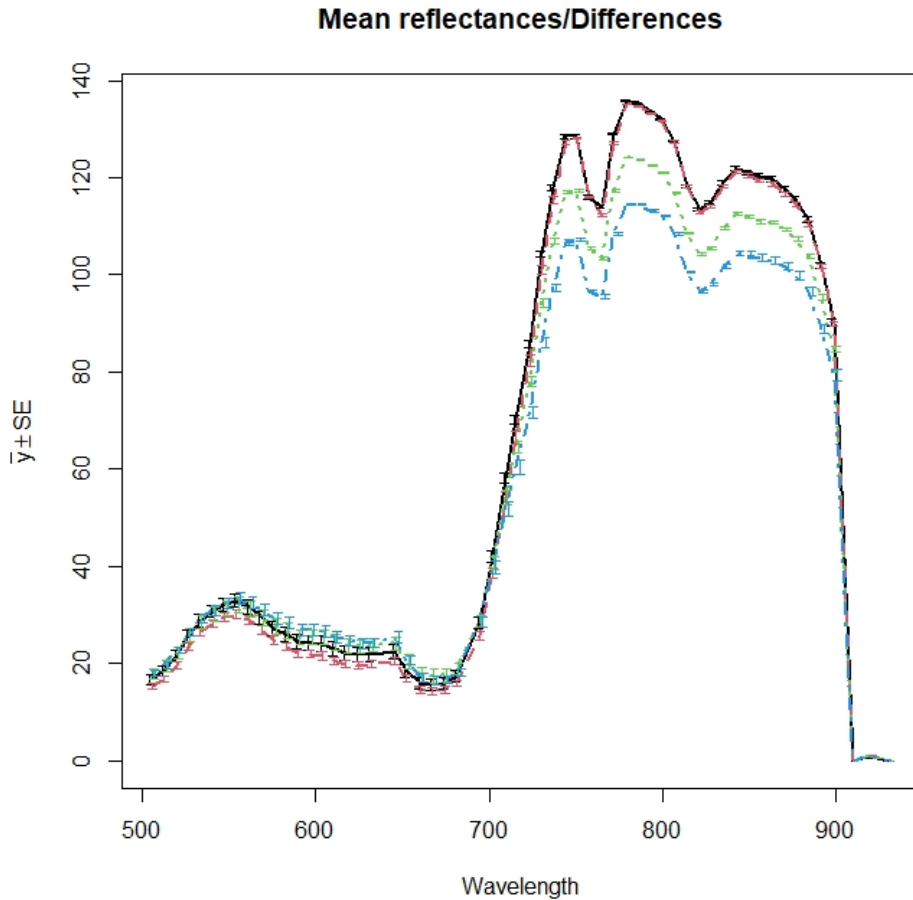
SVM:lla on monia erilaisia käyttötapoja caret-paketissa. Tässä tutkimuksessa käytettiin epälineaaria polynomeihin perustuvia muunnosfunktioita. Polynomeihin perustuvien ytimien käytön etuna on, että se tekee piirteistykseen lähtöavaruudessa. Mallin säädettäviin hyperparametreihin kuuluu C (engl. cost), joka määrittää virheelliset luokitukset ikään kuin sakottaen luokittelijaa niistä (RPubs n.d.). Mitä korkeampi C:n arvo sitä todennäköisempää on, että luokittelutulokset ovat luotettavia. Lisäksi mallissa voi säätää skaalaa (engl. scale) ja polynomitermien maksimaalista astetta (engl. degree). Skaalauksella malli säätää piirteiden arvot samalle vaihteluvälille ja aste puolestaan ilmaisee aineiston muuntamisessa käytetyn polynomin astetta, joka siis määrittää jakamiseen käytettyä funktiota. Luokittelu tehtiin tässä tutkimuksessa mallin oletuksena olleilla hyperparametrien arvoskaaloilla.

4 TULOKSET

Puulajikohtaiset pikseleiden heijastavuuksien keskiarvot eri aallonpituuksilla esikäsittelyiden jälkeen on esitetty Kuvassa 4. Kuvaan on lisäksi merkitty pikseleiden heijastavuuksissa esiintynyt hajonta eri aallonpituuksilla. Kuvasta voidaan havaita heijastavuuksien olleen samankaltainen kaikilla puulajeilla vihreän valon aallonpituuksilla, kunnes lajit alkoivat erottua hieman toisistaan

punaisen valon aallonpituusalueelle siirryttäessä. Parhaiten spektrien väliset erot tulivat esiin lähi-infrapun aallonpituuksilla (n. 700–900 nm). Heijastavuudet olivat korkeampia lehti- kuin havupuulajeilla. Lehtipuulajien heijastavuudet olivat hyvin toisiaan vastaavia koko spektrillä mutta niiden väliset erot tulivat puolestaan parhaiten esiin näkyvän valon puolella, eli aallonpituusvälillä 500–700 nm. Havupuiden spektrit erottuivat selkeimmin toisistaan lähi-infrapun aallonpituuksilla, joilla niiden heijastavuusspektreissä oli huomattavia eroja. Männyn spektri oli kyseisillä aallonpituusalueilla kuusta lähempänä lehtipuuiden spektrejä niiden erottuessa kuitenkin selkeästi toisistaan. Kuusilla havaittiin kaikista alhaisimmat pikseleiden heijastavuuksien keskiarvot.

Kuvasta 4 voidaan myös todeta eri aallonpituuksilla esiintyneen hajonnan olleen erittäin vähäistä, kun pikseleiden kirkkauserojen tasoitus tehtiin saman puulajin yksilöiden pikseleille. Kuvaan käytettyä aineistoa laskettaessa kirkkauserojen tasoituksessa käytettiin myös kasvillisuusindeksejä, tummimpien pikseleiden hylkäämistä (kvantiili 0,25), kirkkauserojen tasoitusta kaikkien puuyksilöiden pikseleiden välillä. Kasvillisuusindeksien hajonta oli hyvin vähäistä verrattuna heijastavuusarvojen hajontaan, joten kuvasta niiden eroja on vaikea havaita.



Kuva 4. Pikseleiden heijastavuuksien keskiarvospektrit ja eri aallonpituuksilla esiintyvät hajonnat puulajikohtaisesti. Musta spektrikäyrä visualisoi harmaalepän, punainen koivun, vihreä männyn ja sininen kuusen heijastavuuksien keskimääräisiä intensiteettejä ja janat hajontien määriä eri aallonpituuksilla.

Taulukossa 3 on esitetty erilaisilla esikäsittelyillä saatuja tuloksia LDA-, RF- ja SVM-algoritmeille. Askeltavalla LDA:lla havaittiin vain hyvin pieniä eroavaisuuksia kokonaistarkkuuden ja Kappa-arvoissa. Korkein kokonaistarkkuus (98 %, Kappa = 0,96) saatiin käyttämällä esikäsittelyssä tummimpien pikseleiden hylkäämistä, lajikohtaista kirkkauserojen tasoitusta sekä kasvillisuusindeksejä. Lähes yhtä hyvään kokonaistarkkuuteen (97 %, Kappa = 0,96) päästiin muuten samoilla esikäsittelyillä mutta käyttämällä kasvillisuusindeksien sijasta peräkkäisten aallonpituuksien heijastavuuksien erotuksia. Askeltavalla LDA:lla saatiin parhaimmat luokittelun kokonaistarkkuudet verrattaessa kaikkia kolmea luokittelijaa keskenään.

Taulukko 3. Eri esikäsittelymenetelmillä saadut kokonaistarkkuudet ja Kappa-arvot LDA-, RF- ja SVM-luokittelijoille. Drop.dark tarkoittaa 25 prosentin tummimmista pikseleistä hylkäämistä, Adjust tasoittaa puuyksilön pikseleiden väliset kirkkauserot, Adj.species tekee saman kaikille saman puulajin pikseleille, Add.deriv käyttää peräkkäisten aallonpituuksien heijastavuuksien erotuksia ja Add.vis kaikkia neljää kasvillisuusindeksiä.

Esikäsittelymenetelmä	Kokonaistarkkuus/Kappa		
	LDA	RF	SVM
Drop.dark, Adjust, Adj.species	0,94/ 0,91	0,95/ 0,94	0,85/0,80
Drop.dark, Adjust, Adj.species, Add.deriv	0,95/0,93	0,91/0,89	0,87/0,82
Drop.dark, Adjust, Adj.species, Add.vis	0,95/0,93	0,95/0,93	0,86/0,80
Drop.dark, Adjust, Adj.species, Add.deriv, Add.vis	0,95/0,93	0,93/0,91	0,88/0,83
Drop.dark, Adj.species, Add.vis	0,98/0,96	0,94/0,92	0,78/0,71
Drop.dark, Add.deriv, Adj.species	0,97/0,96	0,94/0,91	0,86/0,81

RF:lla kaikki esikäsittelymenetelmät tuottivat myös yli 90 % luokittelun kokonaistarkkuuden. Korkein saavuttu kokonaistarkkuus oli 95 % (Kappa = 0,93–0,94) ja siihen päästiin kahdella erilaisella esikäsittelymenetelmällä. Ensimmäisessä niistä käytettiin puuyksilöiden tummimpien pikseleiden hylkäämistä, niiden pikseleiden kirkkauserojen tasoitusta ja saman lajin yksilöiden kaikille pikseleille tehtyä kirkkauserojen tasoitusta. Toisessa tehtiin samat vaiheet mutta lisänä käytettiin kasvillisuusindeksejä. Tästä voidaan päätellä, etteivät kasvillisuusindeksit vaikuttaneet luokittelutulokseen.

Kaikissa Taulukossa 3 näkyvissä ajoissa SVM-algoritmin skaalahyperparametrin arvo oli 0,100 ja niissä käytettiin joko ensimmäisen tai toisen asteen polynomia niiden tuottaessa korkeimmat kokonaisluokittelutarkkuudet ja Kappa-arvot. SVM:lla korkeimmaksi luokittelun kokonaistarkkuudeksi saatiin 88 % (Kappa = 0,83), mikä oli kuitenkin selkeästi matalampi kuin kahdella muulla luokittelijalla. Tällöin sakkoarvo eli C oli 0,25. Yleisesti ottaen SVM:lla saaduissa tuloksissa oli myös eniten vaihtelua erilaisten esikäsittelymenetelmien välillä. Alhaisimmillaan luokittelun kokonaistarkkuus oli 78 % (Kappa = 0,71) käytettäessä tummimpien pikseleiden hylkäämistä, lajikohtaista kirkkauserojen tasoitusta sekä kasvillisuusindeksejä, mikä puolestaan tuotti LDA:lla parhaan luokittelun kokonaistarkkuuden ja Kappa-arvon koko tutkimuksessa. Muilla esikäsittelyillä SVM:n tulokset olivat yli 84 % (Kappa = 0,80–0,83).

5 TULOSTEN TARKASTELU

Askeltavalla LDA:lla saadut korkeimmat luokittelun kokonaistarkkuudet (97–98 %) päihittivät luokittelijoita vertailtaessa RF:lla ja SVM:lla saadut tulokset. LDA-algoritmin ei oletettu toimivan tämän aineiston kohdalla hyvin, sillä se olettaa piirteiden välistä riippumattomuutta. Tässä aineistossa aallonpituuskaistojen heijastavuuksien välillä oli korrelaatiota. Tavallista LDA-luokittelijaa käytettäessä R antoikin varoituksen muuttujien kollineaarisuudesta. Mitä vähemmän kollineaarisuutta selittävien muuttujien välillä havaitaan sitä itsenäisempiä ne ovat selitettävän muuttujan käyttäytymisen selittämisessä. Tästä huolimatta askeltavaa muuttujien valintaa käytettäessä LDA tuotti parhaat kokonaistarkkuudet tutkimuksessa mukana olleita neljää puulajia (harmaaleppä, koivu, mänty ja kuusi) luokiteltaessa. Syynä tähän on varmaankin vertailuajoissa käytetty askeltava muuttujien valinta, joka jätti automaattisesti korreloivia piirteitä pois. Tälle aineistolle LDA-luokittelijan toimintaperiaate vaikutti sopivan, ja malli suoriutui luokkien välisen erottuvuuden maksimoinnissa hyvin.

PCA-analyysi ja askeltava LDA-analyysi ovat siitä samankaltaisia, että molemmat pyrkivät kaventamaan lähtöaineiston moniulotteisuutta. Tämän aineiston tapauksessa PCA ei kuitenkaan parantanut luokittelutuloksia. Varmaa syytä tähän on mahdotonta antaa mutta todennäköisesti käytetyssä aineistossa oli liian paljon muuttujia suhteessa näytteiden määrään. Tulevissa tutkimuksissa PCA-analyysin toimivuutta voidaan oletettavasti parantaa kasvattamalla koulutuksessa ja luokittelussa käytettävien näytteiden lukumäärää. Tässä tutkimuksessa aineistoa jouduttiin muutenkin keräämään useaan otteeseen, ja kokonaan uuden aineiston kerääminen olisi pitkittänyt tutkimuksen valmistumista, minkä vuoksi päätettiin lopulta turvautua tutkimuksen ulkopuoliseen aineistoon. Verrattaessa PCA- ja LDA-algoritmien oletuksia teoriatasolla, olisi lähtökohtaisesti voitu olettaa PCA-analyysin parantavan tämän tutkimuksen aineiston luokittelutuloksia, sillä näytteiden vähäinen määrä ei vaikuta siihen tyypillisesti yhtä paljon kuin ohjatun oppimisen luokittelualgoritmeihin.

RF-algoritmin soveltuvuuden tämän tutkimuksen luokittelutehtävään todettiin olevan hyvä, mitä tukevat myös aiempien puulajinluokittelusta tehtyjen tutkimusten tulokset. Tässä tutkimuksessa saavutetut luokittelun kokonaistarkkuudet olivatkin samansuuntaisia tai jopa osin hieman korkeampia kuin taustaosiossa esiteltyjen tutkimusten tulokset. On kuitenkin syytä huomioida, että tutkimusten keskinäinen vertailtavuus on alhainen toisistaan poikkeavien toteutustapojen vuoksi. Yksi silmiinpistävimmistä tuloksista tässä tutkimuksessa oli RF-algoritmin suoriutuminen tasaisemmin ja luokittelun kokonaistarkkuuksiltaan paremmin kuin SVM-algoritmi. Lukuisissa tutkimuksissa (mm. Ballantin ym. 2016, Cao ym. 2020, Viinikka 2020) SVM-luokittelijan avulla on päästy korkeampiin luokittelun kokonaistarkkuuksiin kuin RF-luokittelijalla. Tähän voi olla monia syitä mutta kaikista merkittävimiksi vaikuttajiksi arvioisin käytetyn SVM-menetelmän valinnan sekä hyperparametrien säätämisen puutteellisuuden luokitellulle aineistolle sopivaksi. Tämän tutkimuksen aikataulun puitteissa laajamittaista menetelmien ja hyperparametrien toimivuuden testaamista ei kyetty toteuttamaan mutta on mahdollista, että parempiin luokittelun kokonaistarkkuuksiin olisi tällä tavoin päästy. Toinen syy luokittelijan tuottamiin

alhaisempiin kokonaistarkkuuksiin voisi olla moniluokkainen aineisto. Kuten Melgani ja Bruzzone (2004) kirjoittivat, SVM on alun perin kaksiluokkaiseen luokitteluun suunniteltu malli, joka ei välttämättä tuota yhtä tarkkoja tuloksia moniluokkaisessa luokittelussa. Kaiken kaikkiaan SVM on aiemmin toiminut puulajien luokittelussa hyvin ja sillä saatiin tässäkin tutkimuksessa pääosin hyviä luokittelun kokonaistarkkuuksia, joten pidän ensimmäiseksi mainittua todennäköisempänä syynä mallin alempaan suoriutumiseen kahteen muuhun luokittelijaan verrattuna.

Tämän tutkimuksen tavoitteena oli selvittää, 1) voidaanko 55-taajuuskaistaisista hyperspektrikuviin erottaa mänty, kuusi, harmaaleppä ja koivu toisistaan vähintään 75 % varmuudella, 2) mitkä tutkimukseen valituista esikäsittely- ja luokittelumenetelmistä toimivat parhaiten hankitulle hyperspektrikuva-aineistolle ja 3) auttavatko aineistosta laskettavissa olevat piirteet, kuten peräkkäisten taajuuskaistojen heijastavuuksien erotukset ja kasvillisuusindeksit, puulajien tunnistuksessa. Ensimmäiseen tutkimuskysymykseen vastaus oli, että tutkimuksen aineistolla ja sen analysointiin käytetyillä esikäsittely- ja luokittelumenetelmillä puulajit olivat tunnistettavissa toisistaan yli 77 % kokonaistarkkuudella ja eri luokittelijoilla saaduista tuloksista laskettujen keskiarvojen perusteella noin 90 % tarkkuudella esikäsittelymenetelmästä riippuen. Tulokset viittaavat boreaalisista metsistä kerättyjen hyperspektrikuva-aineistojen soveltuvan luokiteltavaksi ohjatun koneoppimisen menetelmillä ja tuottavan kohtuullisen hyviä ja suurimmalta osin jopa erittäin hyviä kokonaistarkkuuden tuloksia ja Kappa-arvoja.

Toiseen tutkimuskysymykseen vastaus oli tämän aineiston tapauksessa yllättävä; askeltavalla LDA:lla saavutettiin paras luokittelun kokonaistarkkuus (98 %, Kappa = 0,96) käyttämällä esikäsittelyssä tummimpien pikseleiden hylkäämistä, saman puulajin pikseleiden kirkkauserojen tasoitusta sekä kasvillisuusindeksejä. Tämä esikäsittely toimi hyvin myös RF:lla (kokonaistarkkuus = 94 %, Kappa = 0,92). Sen sijaan SVM-algoritmilla kyseinen menetelmä tuotti heikoimman luokittelun kokonaistarkkuuden (78 %, Kappa = 0,71). Samalla esikäsittelyllä, johon lisättiin puuyksilökohtainen valoisuuserojen tasoitus (Adjust) SVM pääsi kuitenkin 86 %

kokonaistarkkuuteen (Kappa = 0,80), mistä voidaan päätellä, että kyseinen esikäsitteily oli tarpeellinen SVM:lla tämän yhdistelmän tapauksessa. Yleisesti ottaen SVM tuotti heikompia luokittelun kokonaistarkkuuden arvoja kuin LDA ja RF, joilla saadut tulokset eivät poikenneet huomattavasti toisistaan. Laskettaessa kaikkien kolmen luokittelijan kokonaistarkkuuksien keskiarvo erilaisilla esikäsitteilymenetelmillä, toimivin esikäsitteilymenetelmien yhdistelmä oli tummimpien pikseleiden hylkääminen (Drop.dark), saman lajin pikseleiden kirkkauserojen tasoitus (Adj.species) ja peräkkäisten aallonpituuksien heijastavuuksien erotukset (Add.deriv), jolloin saavutettu luokittelun kokonaistarkkuuksien keskiarvo oli yli 92 %.

Kolmanteen tutkimuskysymykseen vastaukseksi saatiin, että osalla luokittelijoista kasvillisuusindeksit ja peräkkäisten aallonpituuksien heijastavuuksien erotukset nostivat ja osalla taas laskivat luokittelun kokonaistarkkuuksia hieman. SVM-algoritmillä yhdistettäessä peräkkäisten aallonpituuksien heijastavuuksien erotukset tai ne ja kasvillisuusindeksit muihin esikäsitteilymenetelmiin (Drop.dark, Adjust, Adj.species) saavutettiin 1-3 %-yksikköä korkeampi luokittelun kokonaistarkkuus kuin käyttämällä muita edellä mainittuja esikäsitteilymenetelmiä samanaikaisesti. RF:llä puolestaan lasketuilla piirteillä ei ollut vaikutusta luokittelun kokonaistarkkuuteen tai niiden käyttö heikensi luokittelun kokonaistarkkuuden arvoa hieman (1-4 %-yksikköä). Askeltavalla LDA:lla peräkkäisten aallonpituuksien heijastavuuksien erotukset ja kasvillisuusindeksit yhdessä tai erikseen kasvattivat luokittelun kokonaistarkkuuksia hieman (1-4 %-yksikköä) riippuen esikäsitteilymenetelmien yhdistelmästä. Tässä asiassa ei siis havaittu kaikkiin luokittelijoihin yleispätevää trendiä ja laskettujen piirteiden vaikutuksia luokittelun tarkkuuteen voidaankin pitää riippuvaisena muista esikäsitteilyyn käytetyistä menetelmistä ja luokittelumenetelmästä.

Tässä tutkimuksessa saadut tulokset vahvistavat käsitystä siitä, että puulajien luokittelu hyperspektrikuvista on potentiaalinen menetelmä puulajien tunnistukseen. Tuloksista voidaan kuitenkin todeta, että luokittelun tarkkuutta ohjailee vahvasti kerätyn aineiston ominaisuudet ja laatu sekä sen esikäsitteilyyn ja

luokitteluun valitut menetelmät. Tämän tutkimuksen tulokset tukevat muissa tutkimuksissa saatuja lupaavia tuloksia ohjatun koneoppimisen hyödyntämisestä muiden puulajintunnistukseen käytettävien menetelmien apuna. Hyperspektrikuvauksella voisi esimerkiksi olla mahdollista paikata laserkeilauksessa ja ilmakuvauskuksessa tulevia virheitä puulajinluokittelussa. Tulevaisuudessa tähän tutkimukseen valittujen puulajien luokittelua voitaisiin tutkia lisää monimutkaisemmilla koneoppimisen menetelmillä, kuten (syväoppimiseen tarkoitettuilla) neuroverkoilla. Niiden on todettu optimaalisesti säädettynä tuottavan erittäin hyviä luokittelutarkkuuksia. Myös muunlaisten kasvillisuusindeksien hyödyllisyyttä esikäsittelyssä voisi olla mielenkiintoista testata. Lisäksi SVM-luokittelijan optimointi puulajiluokitteluun olisi yksi mahdollinen tulevaisuuden tutkimusaihe.

Vaikeasti ennustettavat ja globaalit ympäristöongelmat, kuten ilmastonmuutos, tulevat lisäämään tehokkaiden tutkimusmenetelmien tarvetta metsäluonnon monimuotoisuuden tutkimuksessa. Samalla, kun maastotyön tekeminen on muuttunut kalliimmaksi, kansallinen ja kansainvälinen metsiensuojelun lainsäädäntö tulevat kiristymään, jolloin aineiston keruuta ja käsittelyä tehostavien menetelmien merkitys korostuu entisestään. Hyperspektrikuvauksen avulla pystytään tuottamaan erittäin tarkkoja ja informatiivisia aineistoja, joille löytyy lukemattomia sovelluskohteita metsien taloudellisen arvon ja suojeluarvojen kasvattamista ajatellen. Hyperspektrikuvauksen toteuttaminen vaatii alkuinvestointeja mutta on kuitenkin kokonaisuudessaan jo suhteellisen kustannustehokasta. Tutkimus metsän puulajisuhteiden määrittämisestä hyperspektrikuvauksen avulla on kuitenkin vasta lähtökuopissa, sillä etenkin tiheistä, nuorista ja monimuotoisista metsistä saatujen hyperspektrikuva-aineistojen laadullisen vaihtelun hallinta edellyttää pitkäjänteistä ja tavoitteellista kehitystyötä menetelmätasolla.

KIITOKSET

Haluun esittää lämpimän kiitoksen työn valmistumista edistäneestä tuesta ja korvaamattoman tärkeistä ohjeista ohjaajilleni Anssi Lensulle ja Pauliina Salmelle. Kiitokset Anssi Lensulle myös hänen teknisen osaamisensa jakamisesta R-koodien laatimisen ja kehittämisen muodossa. Aineiston esikäsittelyssä ja luokittelussa käytetyt R-skriptit olivat hänen käsialaansa.

KIRJALLISUUS

- Aasen H. & Bolten A. 2018. Multi-temporal high-resolution imaging spectroscopy with hyperspectral 2D imagers – From theory to application. *Remote sensing of environment* 205: 374–389.
- Aasen H., Burkart A., Bolten A. & Bareth G. 2015. Generating 3D hyperspectral information with lightweight UAV snapshot cameras for vegetation monitoring: From camera calibration to quality assurance. *ISPRS Journal of Photogrammetry and Remote Sensing* 108: 245–259, doi:10.1016/j.isprsjprs.2015.08.002.
- Alcolea A., Paoletti M.E., Haut J.M., Resano J. & Plaza A. 2020. Inference in Supervised Spectral Classifiers for On-Board Hyperspectral Imaging: An Overview. *Remote sensing* 12: 534.
- Ballanti L., Blesius L., Hines E. & Kruse B. 2016. Tree Species Classification Using Hyperspectral Imagery: A Comparison of Two Classifiers. *Remote sensing (Basel, Switzerland)* 8: 445, doi:10.3390/rs8060445.
- Breiman, L. 2001. Random forests. *Machine Learning* 45: 5–32.
- Cao J., Liu K., Liu L., Zhu Y., Li J & He Z. 2018. Identifying Mangrove Species Using Field Close-Range Snapshot Hyperspectral Imaging and Machine-Learning Techniques. *Remote Sensing* 10: 2047, doi:10.3390/rs10122047.
- Chen Y., Lin Z., Zhao X., Wang G. & Gu Y. 2014. Deep Learning-Based Classification of Hyperspectral Data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 7 (6): 2094–2107.
- Clark M.L. 2020. Comparison of multi-seasonal Landsat 8, Sentinel-2 and hyperspectral images for mapping forest alliances in Northern California. *ISPRS journal of photogrammetry and remote sensing* 159: 26–40.

- Clark M.L., Buck-Diaz J. & Evens J. 2018. Mapping of forest alliances with simulated multi-seasonal hyperspectral satellite imagery. *Remote Sens. Environ.*: 210, 490–507.
- Colomina I. & Molina P. 2014. Unmanned aerial systems for photogrammetry and remote sensing: A review. *ISPRS Journal of Photogrammetry and Remote Sensing* 92: 79–97, doi: 10.1016/j.isprsjprs.2014.02.013.
- Dadon A., Mandelmilch M., Ben-Dor E. & Sheffer E. 2019. Sequential PCA-based Classification of Mediterranean Forest Plants using Airborne Hyperspectral Remote Sensing. *Remote Sensing* 11: 2800, doi:10.3390/rs11232800.
- Eugenio F., Schons C., Mallmann C., Schuh M., Fernandes P. & Badin T. 2020. Remotely piloted aircraft systems and forests: a global state of the art and future challenges. *Canadian journal of forest research* 50: 705–716, doi:10.1139/cjfr-2019-0375.
- European Environment Agency. 2020. *State of nature in Europe: a health check*. <https://www.eea.europa.eu/themes/biodiversity/state-of-nature-in-the-eu> (luettu: 2.2.2021)
- Fricker G., Ventura J., Wolf J., North M., Davis F. & Franklin J. 2019. A Convolutional Neural Network Classifier Identifies Tree Species in Mixed-Conifer Forest from Hyperspectral Imagery. *Remote Sensing* 11: 2326, doi:10.3390/rs11192326.
- Grahn H. F. & Geladi P. 2007. *Techniques and Applications of Hyperspectral Image Analysis*. John Wiley & Sons Ltd., Englanti.
- Hagen N., Kester R. T., Gao L. & Tkaczyk T. S. 2012. Snapshot advantage: a review of the light collection improvement for parallel high-dimensional measurement systems. *Optical Engineering* 51(11): 111702, doi: 10.1117/1.OE.51.11.111702.
- Hastie T. J., Tibshirani R. J., & Friedman J. H. 2009. *The elements of statistical learning*. Springer, Englanti.
- Hennessy A., Clarke K. & Lewis M. 2020. Hyperspectral Classification of Plants: A Review of Waveband Selection Generalisability. *Remote Sensing* 12: 113.
- Hong G. & Zhang Y. 2008. A comparative study on radiometric normalization using high resolution satellite images. *International journal of remote sensing* 29: 425–438.
- Honkavaara E., Eskelinen M.A., Pölönen I., Saari H., Ojanen H., Mannila R., Holmlund C., Hakala T., Litkey P., Rosnell T., Viljanen N. & Pulkkanen M. 2016. Remote Sensing of 3-D Geometry and Surface Moisture of a Peat Production Area Using Hyperspectral Frame Cameras in Visible to Short-Wave Infrared Spectral Ranges Onboard a Small Unmanned Airborne Vehicle (UAV). *IEEE transactions on geoscience and remote sensing* 54: 5440–5454.
- Honkavaara E., Saari H., Kaivosoja J., Pölönen I., Hakala T., Litkey P., Mäkynen J. & Pesonen L. 2013. Processing and Assessment of Spectrometric, Stereoscopic

- Imagery Collected Using a Lightweight UAV Spectral Camera for Precision Agriculture. *Remote sensing* (Basel, Switzerland) 5: 5006–5039.
- Horning N., Robinson J. A., Sterling E. J., Turner W. & Spector S. 2010. *Remote Sensing for Ecology and Conservation*. Oxford University Press Inc., New York.
- Hovi A., Raitio P. & Rautiainen M. 2017. A spectral analysis of 25 boreal tree species. *Silva Fennica* 51(4): 7753, doi:10.14214/sf.7753.
- Jenal A., Bareth G., Bolten A., Kneer C., Weber I. & Bongartz J. 2019. Development of a VNIR/SWIR Multispectral Imaging System for Vegetation Monitoring with Unmanned Aerial Vehicles. *Sensors* 19: 5507, doi:10.3390/s19245507.
- Kiranyaz S., Avci O., Abdeljaber O., Ince T., Gabbouj M. & Inman D. J. 2021. 1D convolutional neural networks and applications: A survey. *Mechanical Systems and Signal Processing* 151: 107398.
- Luke. 2015. Valtakunnan metsien inventointi (VMI). <http://www.metla.fi/ohjelma/vmi/info.htm> (luettu: 2.2.2021)
- Luke. 2021. VMI-tiedon avulla kestävä metsätaloutta. <https://www.luke.fi/blogi/vmi-tiedon-avulla-kestavaa-metsataloutta/> (luettu 27.2.2022)
- Lukeš P., Stenberg P., Rautiainen M., Möttus M. & Vanhatalo K. 2013. Optical properties of leaves and needles for boreal tree species in Europe. *Remote Sensing Letters* 4 (7): 667–676, doi:10.1080/2150704X.2013.782112.
- Mannila R., Holmlund C., Ojanen H., Näsilä A. & Saari H. 2014. Short-Wave Infrared (SWIR) spectral imager based on Fabry–Pérot interferometer for remote sensing. *Proceedings Volume 9241, Sensors, Systems, and Next-Generation Satellites XVIII* 9241: 1–8, doi: 10.1117/12.2067206.
- Markiet V. & Möttus M. 2020. Estimation of boreal forest floor reflectance from airborne hyperspectral data of coniferous forests. *Remote Sensing of Environment* 249: 112018, doi:10.1016/j.rse.2020.112018.
- Melgani F. & Bruzzone L. 2004. Classification of Hyperspectral Remote Sensing Images With Support Vector Machines. *IEEE Transactions on Geoscience and Remote Sensing* 42 (8): 1778–1790, doi: 10.1109/TGRS.2004.831865
- Miyoshi G.T., Arruda M.d.S., Osco L.P., Marcato Junior J., Gonçalves D.N., Imai N.N., Tommaselli A.M.G., Honkavaara E. & Gonçalves W.N. 2020. A Novel Deep Learning Method to Identify Single Tree Species in UAV-Based Hyperspectral Images. *Remote sensing* (Basel, Switzerland) 12: 1294.
- Möttus M., Sulev M. & Hallik L. 2014. Seasonal Course of the Spectral Properties of Alder and Birch Leaves. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 7: 2496–2505.
- Modzelewska A., Fassnacht F.E. & Stereńczak K. 2020. Tree species identification within an extensive forest area with diverse management regimes using airborne hyperspectral data. *ITC journal* 84: 101960, doi:10.1016/j.jag.2019.101960.

- Mozgeris G., Juodkienė V., Jonikavičius D., Straigyte L., Gadal S. & Ouerghemmi W. 2018. Ultra-Light Aircraft-Based Hyperspectral and Colour-Infrared Imaging to Identify Deciduous Tree Species in an Urban Environment. *Remote Sensing* 10: 1668, doi:10.3390/rs10101668.
- Nevalainen O., Honkavaara E., Tuominen S., Viljanen N., Hakala T., Yu X., Hyyppä J., Saari H., Pölönen I., Imai N. & Tommaselli A. 2017. Individual Tree Detection and Classification with UAV-Based Photogrammetric Point Clouds and Hyperspectral Imaging. *Remote Sensing* 9: 185, doi:10.3390/rs9030185.
- Nezami S., Khoramshahi E., Nevalainen O., Pölönen I. & Honkavaara E. 2020. Tree Species Classification of Drone Hyperspectral and RGB Imagery with Deep Learning Convolutional Neural Networks. *Remote sensing (Basel, Switzerland)* 12: 1070.
- Näsi R., Honkavaara E., Tuominen S., Saari H., Pölönen I., Hakala T., Viljanen N., Soukkamäki J., Näkki I., Ojanen H. & Reinikainen J. 2016. UAS based tree species identification using the novel FPI based hyperspectral cameras in visible, NIR and SWIR spectral ranges. *International archives of the photogrammetry, remote sensing and spatial information sciences XLI-B1*: 1143–1148, doi:10.5194/isprs-archives-XLI-B1-1149-2016.
- Pascucci S., Pignatti S., Casa R., Darvishzadeh R. & Huang W. 2020. Special Issue “Hyperspectral Remote Sensing of Agriculture and Vegetation”. *Remote Sensing* 12: 3665, doi:10.3390/rs12213665.
- Pölönen I., Saari H., Kaivosoja J., Honkavaara E. & Pesonen L. 2013. Hyperspectral imaging based biomass and nitrogen content estimations from lightweight UAV. *Remote Sensing for Agriculture, Ecosystems, and Hydrology XV*: 8887, doi:10.1117/12.2028624.
- Raczko E. & Zagajewski B. 2017. Comparison of support vector machine, random forest and neural network classifiers for tree species classification on airborne hyperspectral APEX images. *European Journal of Remote Sensing* 50: 1, 144–154, DOI:10.1080/22797254.2017.1299557.
- Rautiainen M. & Lukeš P. 2015. Spectral contribution of understory to forest reflectance in a boreal site: An analysis of EO-1 Hyperion data. *Remote Sensing of Environment* 171: 98–104, doi:10.1016/j.rse.2015.10.009.
- Rautiainen M., Lukeš P., Homolová L., Hovi A., Pisek J. & Möttus M. 2018. Spectral Properties of Coniferous Forests: A Review of In Situ and Laboratory Measurements. *Remote Sensing* 10: 207, doi:10.3390/rs10020207.
- Rautiainen M. & Stenberg P. 2005. Application of photon recollision probability in coniferous canopy reflectance simulations. *Remote Sensing of Environment* 96: 98–107, doi:10.1016/j.rse.2005.02.009.
- RPubs. n.d. SVM with CARET. <https://rpubs.com/uky994/593668> (luettu 13.1.2022)

- Ruiliang P. 2009. Broadleaf species recognition within situhyperspectral data. *International journal of remote sensing* 30: 2759–2779.
- Saarinen N., Vastaranta M., Näsi R., Rosnell T., Hakala T., Honkavaara E., Wulder M. A., Luoma V., Tommaselli A. M. G., Imai N. N., Ribeiro E. A. W., Guimarães R. B., Holopainen M. & Hyypä J. 2017. Uav-based photogrammetric point clouds and hyperspectral imaging for mapping biodiversity indicators in boreal forests. *The International Archives of the Photogrammetry*: 171–175, doi:10.5194/isprs-archives-XLII-3-W3-171-2017.
- Sothe C., De Almeida C.M., Schimalski M.B., La Rosa, L. E. C, Castro J.D.B., Feitosa R.Q., Dalponte M., Lima C.L., Liesenberg V., Miyoshi G.T. & Tommaselli A.M.G. 2020. Comparative performance of convolutional neural network, weighted and conventional support vector machine and random forest for classifying tree species using hyperspectral and photogrammetric data. *GIScience and remote sensing* 57: 369–394.
- Traficom. 2022. Lennä mahdollisimman turvallisesti – lentoturvallisuus. <https://www.droneinfo.fi/fi/koulutusmateriaali/lenna-mahdollisimman-turvallisesti-lentoturvallisuus> (luettu 27.2.2022)
- Trier ØD., Salberg A., Kermit M., Rudjord Ø, Gobakken T., Næsset E. & Aarsten D. 2018. Tree species classification in Norway from airborne hyperspectral and airborne laser scanning data. *European journal of remote sensing* 51: 336–351.
- Tuominen S., Näsi R., Honkavaara E., Balazs A., Hakala T., Viljanen N., Pölönen I., Saari H. & Reinikainen J. 2017. Tree species recognition in species rich area using uav-borne hyperspectral imagery and stereo-photogrammetric point cloud. *The International Archives of the Photogrammetry*: 185–194, doi:10.5194/isprs-archives-XLII-3-W3-185-2017.
- Turner D., Lucieer A., Malenovský Z., King D. & Robinson S. 2014. Spatial Co-Registration of Ultra-High Resolution Visible, Multispectral and Thermal Images Acquired with a Micro-UAV over Antarctic Moss Beds. *Remote sensing (Basel, Switzerland)* 6: 4003-4024.
- Vakalopoulou M. & Karantzalos K. 2014. Automatic Descriptor-Based Co-Registration of Frame Hyperspectral Data. *Remote sensing* 6: 3409–3426.
- Viinikka A., Hurskainen P., Keski-Saari S., Kivinen S., Tanhuanpää T., Mäyrä J., Poikolainen L., Vihervaara P. & Kumpula T. 2020. Detecting European Aspen (*Populus tremula* L.) in Boreal Forests Using Airborne Hyperspectral and Airborne Laser Scanning Data. *Remote Sensing* 12: 2610, doi:10.3390/rs12162610.
- Wu Y. & Zhang X. 2019. Object-Based Tree Species Classification Using Airborne Hyperspectral Images and LiDAR Data. *Forests* 11: 32, doi:10.3390/f11010032.

- Xue J. & Su B. 2017. Significant Remote Sensing Vegetation Indices: A Review of Developments and Applications. *Journal of Sensors*: doi:10.1155/2017/1353691.
- Yang R. & Kan J. 2020. Classification of tree species at the leaf level based on hyperspectral imaging technology. *Journal of Applied Spectroscopy* 87 (1):184-193, doi:10.1007/s10812-020-00981-9.
- Zhao D., Pang Y., Liu L. & Li Z. 2020. Individual Tree Classification Using Airborne LiDAR and Hyperspectral Data in a Natural Mixed Forest of Northeast China. *Forests* 11: 303, doi:10.3390/f11030303.
- Ymparisto.fi. 2020. Luontodirektiivin luontotyypit. https://www.ymparisto.fi/fi-FI/Luonto/Luontotyypit/Luontodirektiivin_luontotyypit/ (luettu 2.2.2021)